

STIC-Biotech/ChemLib

158263

me

From: Slobodyansky, Elizabeth  
Sent: Tuesday, July 05, 2005 7:58 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/758,672 SEQ

Please search for case 10/758,672:

88

SEQ ID NOs: 1 and 2 against commercial ~~and interference~~ databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner  
Art Unit 1652  
REM 2D83  
571-272-0941  
MAILBOX 2C70

\*\*\*\*\*  
STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*  
Type of Search

NA#: \_\_\_\_\_ AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*  
Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 18:41:46 ; Search time 25839 Seconds  
(without alignments)  
11829.231 Million cell updates/sec

Title: US-10-758-672A-1  
Perfect score: 6308  
Sequence: 1 gccaaagattcgccagcagg.....aattttgtattggtgtttt 6308

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba:\*  
2: gb.hg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.ats:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6308	100.0	6308	6	AR487554
2	5250	83.2	5250	9	AY061886 Homo sapi
3	5136	81.4	5205	6	AR487565 Sequence
4	5085	80.6	5128	9	AP525401 Homo sapi
5	4435.4	70.3	6395	6	AR030784 Sequence
6	4435.4	70.3	6395	6	AR121463 Sequence
7	4435.4	70.3	6395	6	BD245011 Nucleic a
8	4435.4	70.3	6395	10	AF061555 Mus muscu
9	3473.6	55.1	4529	6	CQ728565 Sequence
10	3055.8	48.4	3059	6	AX714232 Sequence
11	3055.8	48.4	3059	9	AK056441 Homo sapi
12	2546.8	40.4	2550	6	BD156870 Primer fo
13	2546.8	40.4	2550	6	AX877826 Sequence
14	2546.8	40.4	2550	9	AK027803 Homo sapi
15	2487.8	39.4	4576	9	BS648400 Homo sapi
16	1529.2	24.2	3598	6	CQ850354 Sequence
17	1529.2	24.2	3598	9	AK127478 Homo sapi
18	1381.8	21.9	6300	6	AR487555 Sequence
19	1378.6	21.9	5268	9	AY061884 Homo sapi

20	1233.6	19.6	6089	6	AR487556
21	1233.6	19.6	7511	10	BC075642
22	1233.4	19.6	5265	10	AY061885
23	1217.4	19.3	5438	10	AY280958
24	996.2	15.8	1001	6	AR030785
25	996.2	15.8	1001	6	AR121464
26	996.2	15.8	1001	6	BD245012
27	994.2	15.8	999	9	AF061556
28	981	15.6	6158	9	AB002347
29	870.2	13.8	6521	10	AK122254
30	858	13.6	1330	5	CR387019
31	813.6	12.9	3327	6	CQ755169
32	813.6	12.9	3502	6	AX965629
33	797.2	12.6	4208	6	CQ842877
34	797.2	12.6	4208	9	AK125795
35	772.8	12.3	3502	10	BC031403
36	764.2	12.1	818	9	HXS521
37	656.8	10.4	733	6	BD221289
38	651.6	10.3	2958	9	AK026998
39	647	10.3	712	6	BD149613
40	647	10.3	712	6	AX869551
41	638.6	10.1	756	6	BD221205
42	515.2	8.2	2475	10	BC026391
43	503.8	8.0	16518	9	AC090514
44	487.8	7.7	2512	9	AK026948
45	392.4	6.2	2518	9	BC064512

## ALIGNMENTS

RESULT 1  
AR487554  
LOCUS AR487554 6308 bp DNA linear PAT 14-MAY-2004  
DEFINITION Sequence 1 from patent US 6706505.  
ACCESSION AR487554  
VERSION AR487554.1 GI:47252804  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 6308)  
AUTHORS Han, H.-Q. and Kwak, K.  
TITLE Human E3.alpha. ubiquitin ligase family  
JOURNAL Patent: US 6706505-A 1 16-MAR-2004;  
FEATURES Location/Qualifiers  
source 1..6308  
/organism="unknown"  
/mol\_type="genomic DNA"

/mol_type= genomic DNA									
ORIGIN	Query Match	100.0%;	Score 6308;	DB 6;	Length 6308;				
	Best Local Similarity	100.0%;	Pred. No. 0;						
	Matches 6308;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
Qy	1	GCCAAAGATTTCGGCAGCGGGGAAAAGCTGAGCCCGAGGAAACCAAATTACTTGTGCTTTACCT	60						
Db	1	GCCAAAGATTTCGGCAGCGGGGAAAAGCTGAGCCCGAGGAAACCAAATTACTTGTGCTTTACCT	60						
Qy	61	CATTGTGTAAAGCAACAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAGAACTTTACCT	120						
Db	61	CATTGTGTAAAGCAACAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAGAACTTTACCT	120						
Qy	121	CCAAAGGCTTATCATCTGTCTTCCACTTATCCAAACAGCTGTATGGCCACTGCGCTGTGC	180						
Db	121	CCAAAGGCTTATCATCTGTCTTCCACTTATCCAAACAGCTGTATGGCCACTGCGCTGTGC	180						
Qy	181	CGCACCTTGGAAACCAACGCGCAGGCCCACTACTGCTCCACTACCACTGGTGTCTCCCAACCT	240						
Db	181	CGCACCTTGGAAACCAACGCGCAGGCCCACTACTGCTCCACTACCACTGGTGTCTCCCAACCT	240						
Qy	241	GATCAGCTGCTTGTCTGCTGCCATCTTATCCGCTTCTGCGCTGTCTCGAGTAATAATGTTATACA	300						

Db 241 GATCAGCTGCTTGCTGCCATCTTATCGCTTCTGAGTAAATGTATACA 300  
Qy  
301 CACCTTGAACACCAATCTTACTTTCTGTGCTATGAATTTGACTACTCTAGCTGGATC 360  
Db 301 CACCTTGAACACCAATCTTACTTTCTGTGCTATGAATTTGACTACTCTAGCTGGATC 360  
Qy 361 CGAGCTTTTGTATACATGTCGAAGTGGCCACGGGGTAGAATCCTTAAATAATAGAGAT 420  
Db 361 CGAGCTTTTGTATACATGTCGAAGTGGCCACGGGGTAGAATCCTTAAATAATAGAGAT 420  
Qy 421 GTATGCAACAGTTCACGACCAAAACCCAGATATACAAACCTTACAGCTACCAAGAGTAC 480  
Db 421 GTATGCAACAGTTCACGACCAAAACCCAGATATACAAACCTTACAGCTACCAAGAGTAC 480  
Qy 481 GCCTGATTAATTAGAGGGGAAAAAATCTCCAGTCCCTTCAAGCTGTCGCTTGCT 540  
Db 481 GCCTGATTAATTAGAGGGGAAAAAATCTCCAGTCCCTTCAAGCTGTCGCTTGCT 540  
Qy 541 TCCGGGAAGCGGCGGAAGCACTCTCGAGTCTCGCTCAAAACCCGACTTTCAGGGGCG 600  
Db 541 TCCGGGAAGCGGCGGAAGCACTCTCGAGTCTCGCTCAAAACCCGACTTTCAGGGGCG 600  
Qy 601 TCGTAAAAAGTGTGCTCCCTGTCTCTCCGACCGGCCACAGGTTTCGCTCTGGCCG 660  
Db 601 TCGTAAAAAGTGTGCTCCCTGTCTCTCCGACCGGCCACAGGTTTCGCTCTGGCCG 660  
Qy 661 GGGGTGCGGAACCTGACGGGCTCAGTTCCCTCAAGATGGCGGACGAGGAGCTGGAGTA 720  
Db 661 GGGGTGCGGAACCTGACGGGCTCAGTTCCCTCAAGATGGCGGACGAGGAGCTGGAGTA 720  
Qy 721 CTGACAGATGGAATCAGCGCGAGTTACCCAGACCCCTCAGCGTCTGGCATCTGGT 780  
Db 721 CTGACAGATGGAATCAGCGCGAGTTACCCAGACCCCTCAGCGTCTGGCATCTGGT 780  
Qy 781 GGGATCAGCAAGTGAATTTTATCTGCTTTCTGCAATCATTTGGCACAATTTGGTGGCCAG 840  
Db 781 GGGATCAGCAAGTGAATTTTATCTGCTTTCTGCAATCATTTGGCACAATTTGGTGGCCAG 840  
Qy 841 AAATTTACTTGTGCTGAATGGAACCCAGACTTGGAAAAAGCAGGAGAAAGTGTACAAATGT 900  
Db 841 AAATTTACTTGTGCTGAATGGAACCCAGACTTGGAAAAAGCAGGAGAAAGTGTACAAATGT 900  
Qy 901 CAATATTCACCTCACTGGAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960  
Db 901 CAATATTCACCTCACTGGAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960  
Qy 961 AATTGAAGCAGTGGAGCAATTCAGCTTTTGGGAGGGTTTCAAAAAGTGGAGAGACAA 1020  
Db 961 AATTGAAGCAGTGGAGCAATTCAGCTTTTGGGAGGGTTTCAAAAAGTGGAGAGACAA 1020  
Qy 1021 CCTATTCTTGAGGGATTTGCAATTGATCCAAATGTGATCTGTATGGACTGCTTCC 1080  
Db 1021 CCTATTCTTGAGGGATTTGCAATTGATCCAAATGTGATCTGTATGGACTGCTTCC 1080  
Qy 1081 AGGACAGTGTTCATAAAAATCATGTTACAGATGCATCTTCTACTGGAGGAGGTTCT 1140  
Db 1081 AGGACAGTGTTCATAAAAATCATGTTACAGATGCATCTTCTACTGGAGGAGGTTCT 1140  
Qy 1141 GTGACTGTGGAGACACAGAGGCAATGGAAAACTGGCCCTTTTGTGTAAATCATGAACTG 1200  
Db 1141 GTGACTGTGGAGACACAGAGGCAATGGAAAACTGGCCCTTTTGTGTAAATCATGAACTG 1200  
Qy 1201 GAAGAGCGGTACTATATAAGAGAAATTCACGCTGTCGGTTGAATGAAGAGTAAATGTGCC 1260  
Db 1201 GAAGAGCGGTACTATATAAGAGAAATTCACGCTGTCGGTTGAATGAAGAGTAAATGTGCC 1260  
Qy 1261 AAGCAGGAAATATTTCTTCTCAGTGATAAATAATGTCTAGAAATGACTATATGGGAAG 1320  
Db 1261 AAGCAGGAAATATTTCTTCTCAGTGATAAATAATGTCTAGAAATGACTATATGGGAAG 1320  
Qy 1321 AGGAAAAAGAACTGCTCTCGAACTCCAGATAAGGGAGAAAAATGAAAGATACTATTGTG 1380  
Db 1321 AGGAAAAAGAACTGCTCTCGAACTCCAGATAAGGGAGAAAAATGAAAGATACTATTGTG 1380

Qy 1381 TCCCTTTCAATGATGAACACCAATTCATATGACACAGCTCATATACAGCTCAAAAGAGCTC 1440  
Db 1381 TCCCTTTCAATGATGAACACCAATTCATATGACACAGCTCATATACAGCTCAAAAGAGCTC 1440  
Qy 1441 TTGACTGTGAGCTCGCAGAGGCCAGTTGTCATACCTGCTCCATTTGACAAAGAGGGTGTGTC 1500  
Db 1441 TTGACTGTGAGCTCGCAGAGGCCAGTTGTCATACCTGCTCCATTTGACAAAGAGGGTGTGTC 1500  
Qy 1501 GGGCTGTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGGAAGATATAAGAGTTC 1560  
Db 1501 GGGCTGTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGGAAGATATAAGAGTTC 1560  
Qy 1561 ATTCAAGAAAATGTCTCTCAACATCACTCATGTAGGAAGTATTACACTCAGAGATTTATGG 1620  
Db 1561 ATTCAAGAAAATGTCTCTCAACATCACTCATGTAGGAAGTATTACACTCAGAGATTTATGG 1620  
Qy 1621 CTCAATCAGAAAATTTGCTTTGCTGCTTCTGGATGAAACAAAATTTATGAGCTATTCAA 1680  
Db 1621 CTCAATCAGAAAATTTGCTTTGCTGCTTCTGGATGAAACAAAATTTATGAGCTATTCAA 1680  
Qy 1681 GTGACTTTAGGCAGATCTTTTGGCCAGCATGCTTTAGAGAAACCTGACTCGGAGATC 1740  
Db 1681 GTGACTTTAGGCAGATCTTTTGGCCAGCATGCTTTAGAGAAACCTGACTCGGAGATC 1740  
Qy 1741 CCTGCTCATAAAGCAGGTTAATGCTTTTGGGATGCAAGCTTTATAAAAGTCCCGTAAGA 1800  
Db 1741 CCTGCTCATAAAGCAGGTTAATGCTTTTGGGATGCAAGCTTTATAAAAGTCCCGTAAGA 1800  
Qy 1801 TCCCTCATGAATTTGATCTTTCAGCAAGTTTATTTATGAGATGGAATACAAAACCTCTTTG 1860  
Db 1801 TCCCTCATGAATTTGATCTTTCAGCAAGTTTATTTATGAGATGGAATACAAAACCTCTTTG 1860  
Qy 1861 CTATGGAATTTGCTGAAGTATTATAACAACTGCGAGAAAGATATATCAGTGATGATCATG 1920  
Db 1861 CTATGGAATTTGCTGAAGTATTATAACAACTGCGAGAAAGATATATCAGTGATGATCATG 1920  
Qy 1921 ACAGAGTATCTCTATTAACCTGCTTTCAGTTCAGATGTTTCTGTTCTTACTCTGGCTC 1980  
Db 1921 ACAGAGTATCTCTATTAACCTGCTTTCAGTTCAGATGTTTCTGTTCTTACTCTGGCTC 1980  
Qy 1981 GACATCTTTATGAAGAGCAGAAATGTTATCTCTGTCAATTTCTGTAAGAGTTT 2040  
Db 1981 GACATCTTTATGAAGAGCAGAAATGTTATCTCTGTCAATTTCTGTAAGAGTTT 2040  
Qy 2041 TACCTGAGTACTTGGACAGGAACTAATTAATTTCACTTCCAGGGTTATAGCCAGGACAAAT 2100  
Db 2041 TACCTGAGTACTTGGACAGGAACTAATTAATTTCACTTCCAGGGTTATAGCCAGGACAAAT 2100  
Qy 2101 TGGCAAGAGTATATGTCAGTAAATGTCACCTTAAAGTATATCTCTGATCAGCAAAACCCCAA 2160  
Db 2101 TGGCAAGAGTATATGTCAGTAAATGTCACCTTAAAGTATATCTCTGATCAGCAAAACCCCAA 2160  
Qy 2161 TATGGACAGAAAGTATGAAGATGTCAGTTCCTTGAAGGTTTTCGATCTTTTGTGAAGATTC 2220  
Db 2161 TATGGACAGAAAGTATGAAGATGTCAGTTCCTTGAAGGTTTTCGATCTTTTGTGAAGATTC 2220  
Qy 2221 TTAATCTGTATGCAAGGAAATGGAAGAAATCCGAAAGCAGGTTGGGCAACACATTGAAGTGG 2280  
Db 2221 TTAATCTGTATGCAAGGAAATGGAAGAAATCCGAAAGCAGGTTGGGCAACACATTGAAGTGG 2280  
Qy 2281 ATCTCTGATTTGGGAGCTGCCATTCCTATACAGATGCAATTTGAAGAAATATTTTCTCATGT 2340  
Db 2281 ATCTCTGATTTGGGAGCTGCCATTCCTATACAGATGCAATTTGAAGAAATATTTTCTCATGT 2340  
Qy 2341 TCCAAAGAGTGTGTGCTTGTGATGAAGAACTCTTACTTTGTGGCTTATAAAGAAATGTCA 2400  
Db 2341 TCCAAAGAGTGTGTGCTTGTGATGAAGAACTCTTACTTTGTGGCTTATAAAGAAATGTCA 2400  
Qy 2401 AAGCTGTGATGAGGTGCGATCCAGTTCATCTATCTAGTAGCAACAGTAGTACATCGT 2460  
Db 2401 AAGCTGTGATGAGGTGCGATCCAGTTCATCTATCTAGTAGCAACAGTAGTACATCGT 2460



QY 2461 GTGGACATAGTTTGGAAACAAAGTCTTACAGATATCTGAGGATCTTGTAGCATACATC 2520  
Db 2461 GTGGACATAGTTTGGAAACAAAGTCTTACAGATATCTGAGGATCTTGTAGCATACATC 2520  
QY 2521 TGCCACTCTCTAGGACCCCTGCTGCTCTTACATGATCTTAAAGCAGGCTGGTCTGCTTT 2580  
Db 2521 TGCCACTCTCTAGGACCCCTGCTGCTCTTACATGATCTTAAAGCAGGCTGGTCTGCTTT 2580  
QY 2581 CAAGACTGCATGAATTTGCTCTTTTGAAGACTTTCAAGTAGAGTACTAGTGGAAATATC 2640  
Db 2581 CAAGACTGCATGAATTTGCTCTTTTGAAGACTTTCAAGTAGAGTACTAGTGGAAATATC 2640  
QY 2641 CTTTACGTTGCTGGTGTGGTCCAGGTGTTGCTGAGATGTTGGCGAAGAAATGAC 2700  
Db 2641 CTTTACGTTGCTGGTGTGGTCCAGGTGTTGCTGAGATGTTGGCGAAGAAATGAC 2700  
QY 2701 TGTCTCTTATTAGCCAGGTGTTTATTACCAAGATGTTAAAGTGCAGAGAAATGATG 2760  
Db 2701 TGTCTCTTATTAGCCAGGTGTTTATTACCAAGATGTTAAAGTGCAGAGAAATGATG 2760  
QY 2761 ATAAAGATATCATCATGCTTCAGATTTGCTGATCTTTAATGGATCCCAATAGTCTTGT 2820  
Db 2761 ATAAAGATATCATCATGCTTCAGATTTGCTGATCTTTAATGGATCCCAATAGTCTTGT 2820  
QY 2821 TACTGGTACTTCAGAGGTATGAATTCGCGAGGCTTTTAAACAGCCATATCTACAAAG 2880  
Db 2821 TACTGGTACTTCAGAGGTATGAATTCGCGAGGCTTTTAAACAGCCATATCTACAAAG 2880  
QY 2881 ACCAGGATTTGATTAACAAATATATACATTAAGAGAAATGCTTCAGGTCTCATCT 2940  
Db 2881 ACCAGGATTTGATTAACAAATATATACATTAAGAGAAATGCTTCAGGTCTCATCT 2940  
QY 2941 ATATTGTTGGGTGAGGCTTATGCTGAGTGGGAAATGTGACCAAAAGAGGTGCAAA 3000  
Db 2941 ATATTGTTGGGTGAGGCTTATGCTGAGTGGGAAATGTGACCAAAAGAGGTGCAAA 3000  
QY 3001 TGAGAGAAATCAATTCATCTGCTTTGCAATGAAACCCATGCGCACAGTGCCTTACAAA 3060  
Db 3001 TGAGAGAAATCAATTCATCTGCTTTGCAATGAAACCCATGCGCACAGTGCCTTACAAA 3060  
QY 3061 ATTTACCTGAGATGAATAATTAAGAACTGCTGAGAAATGTCAATAACAAAGTGGCCA 3120  
Db 3061 ATTTACCTGAGATGAATAATTAAGAACTGCTGAGAAATGTCAATAACAAAGTGGCCA 3120  
QY 3121 CATTTAAGAAACAGGCTGATCAGGCCATGGAGTTTATGAATTAAGATGAATCACTGA 3180  
Db 3121 CATTTAAGAAACAGGCTGATCAGGCCATGGAGTTTATGAATTAAGATGAATCACTGA 3180  
QY 3181 AAGACTTCAATATGTAATTTTATCAATTAATCAAAACCCAGATAGCAAGGCTGAACATA 3240  
Db 3181 AAGACTTCAATATGTAATTTTATCAATTAATCAAAACCCAGATAGCAAGGCTGAACATA 3240  
QY 3241 TGCAGAAAGAGGAAACAAAGAAACAAAGATGAAGCTGCGCCACCAACCTC 3300  
Db 3241 TGCAGAAAGAGGAAACAAAGAAACAAAGATGAAGCTGCGCCACCAACCTC 3300  
QY 3301 CTGAATTCGCTGCTTTTACAGAAAGTATTAACCTTCTCACTGCTATCATGATGT 3360  
Db 3301 CTGAATTCGCTGCTTTTACAGAAAGTATTAACCTTCTCACTGCTATCATGATGT 3360  
QY 3361 ACATTTCTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCTAATCTTGGACCGAAG 3420  
Db 3361 ACATTTCTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCTAATCTTGGACCGAAG 3420  
QY 3421 GGATGCTCCAAATGGCTTTTATCTTGGCAATGGGTTTACTAGAGAGAGCAACAGC 3480  
Db 3421 GGATGCTCCAAATGGCTTTTATCTTGGCAATGGGTTTACTAGAGAGAGCAACAGC 3480  
QY 3481 TTCAAAAGCTCTTGAAGAGAGTAAATTTGACTTTTATCATTAAGGCTTCAAGATTGG 3540  
Db 3481 TTCAAAAGCTCTTGAAGAGAGTAAATTTGACTTTTATCATTAAGGCTTCAAGATTGG 3540  
QY 3541 GAACTTCAGCCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCGATTAG 3600

Db 3541 GAACTTCAGCCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCGATTAG 3600  
QY 3601 AAGCCAGAGAGCATGATTAACGTGGATCTTTCAGATGTTTTGACACAGTGAAGCGATTAA 3660  
Db 3601 AAGCCAGAGAGCATGATTAACGTGGATCTTTCAGATGTTTTGACACAGTGAAGCGATTAA 3660  
QY 3661 GAGAAAAATCTTGTGTTTAAATTTGTAGCAACCATCAGGATCGGAAATCTTAAAGAAATGATG 3720  
Db 3661 GAGAAAAATCTTGTGTTTAAATTTGTAGCAACCATCAGGATCGGAAATCTTAAAGAAATGATG 3720  
QY 3721 AGATTACTCATGATTAAGAAAAACAGAAACGAAAAAGAAAGCTGAAAGCTCTAGGCTTAC 3780  
Db 3721 AGATTACTCATGATTAAGAAAAACAGAAACGAAAAAGAAAGCTGAAAGCTCTAGGCTTAC 3780  
QY 3781 ATGCCAGAGAGATCATGGCTCAGATGTTGCTCTTACAGAAAAAATCTTCAATGAAACTCAT 3840  
Db 3781 ATGCCAGAGAGATCATGGCTCAGATGTTGCTCTTACAGAAAAAATCTTCAATGAAACTCAT 3840  
QY 3841 AACTCATGATGACATACATCAGAAATGCTCGGAAAGAAAGATTCATTTAGGGAAG 3900  
Db 3841 AACTCATGATGACATACATCAGAAATGCTCGGAAAGAAAGATTCATTTAGGGAAG 3900  
QY 3901 AGACACCCAGCAGTCACTGAGTCTTCTAGAAATTTGCTTGGGCTCTTAAACGGGCTCAT 3960  
Db 3901 AGACACCCAGCAGTCACTGAGTCTTCTAGAAATTTGCTTGGGCTCTTAAACGGGCTCAT 3960  
QY 3961 CTGTTACTGAAAAAGGAGTCTGACGTCATCTTTCGCAAGAAAGAAACAGGAGGTGAAAA 4020  
Db 3961 CTGTTACTGAAAAAGGAGTCTGACGTCATCTTTCGCAAGAAAGAAACAGGAGGTGAAAA 4020  
QY 4021 TAGAAAAATATGCAATGATTTATCGGCTGTGTCAGAAATCTTACGCCCTTAAACCCAGC 4080  
Db 4021 TAGAAAAATATGCAATGATTTATCGGCTGTGTCAGAAATCTTACGCCCTTAAACCCAGC 4080  
QY 4081 ACAGGGAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCATTTTCATGATGTCAG 4140  
Db 4081 ACAGGGAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCATTTTCATGATGTCAG 4140  
QY 4141 ACTTGGCATATGGAATCTTATACAGAAAGCTGTGTCATGTAATGCAAGCTGTGCTGGC 4200  
Db 4141 ACTTGGCATATGGAATCTTATACAGAAAGCTGTGTCATGTAATGCAAGCTGTGCTGGC 4200  
QY 4201 AGAAGTATTTTGAAGCTGTACAGTGTCTCAGAGCGCATTCATGTTGACCTTTTGG 4260  
Db 4201 AGAAGTATTTTGAAGCTGTACAGTGTCTCAGAGCGCATTCATGTTGACCTTTTGG 4260  
QY 4261 ACTTGGAAAGTGGAGAAATCTTGGCTCTTTCGCAATCTCTGTCGAATCTGTGATTC 4320  
Db 4261 ACTTGGAAAGTGGAGAAATCTTGGCTCTTTCGCAATCTCTGTGCAATCTGTGATTC 4320  
QY 4321 CCATTTATTTTTCGCAACCTCAAGATTAACAGTGAAGATGCAAGTCTCTTGTCAAC 4380  
Db 4321 CCATTTATTTTTCGCAACCTCAAGATTAACAGTGAAGATGCAAGTCTCTTGTCAAC 4380  
QY 4381 TTTTGAACCTTGGCAGCGTGCATACAGCTGTTCTGGCCAGAAATATCAGGTTTATAATA 4440  
Db 4381 TTTTGAACCTTGGCAGCGTGCATACAGCTGTTCTGGCCAGAAATATCAGGTTTATAATA 4440  
QY 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTTCTTAAATCAAGGAATGGGAGATTCTA 4500  
Db 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTTCTTAAATCAAGGAATGGGAGATTCTA 4500  
QY 4501 CTTTGGAGTTTCCATTTCCATCTGAGTTTGGGCTTGGAGTCTTCGATTAATAATTTCAATA 4560  
Db 4501 CTTTGGAGTTTCCATTTCCATCTGAGTTTGGGCTTGGAGTCTTCGATTAATAATTTCAATA 4560  
QY 4561 GCATCAAGGAAATGTTTCTTCTTGGCCACAAATTTATAGAAATGGATTGGAAGTGC 4620  
Db 4561 GCATCAAGGAAATGTTTCTTCTTGGCCACAAATTTATAGAAATGGATTGGAAGTGC 4620  
QY 4621 CACCTGATGAAGGAGTCTCGAGTCCCATGCTGACCTGAGGACCTCGGCTTCACTA 4680

Db 4621 CACCTGATGAAGGATCCTCGAGTCCCACTGCTGACCTCGAGCACCTGCGCTTTCACCTA 4680  
Qy TCCAGGCAATTTGAAATCTATTGGGAGATGAAGAAACCTCTCTTTGGAGCACTTCAAA 4740  
Db TCCAGGCAATTTGAAATCTATTGGGAGATGAAGAAACCTCTCTTTGGAGCACTTCAAA 4740  
Qy ATAGGACGATAATAGGCTGGAAGCAATTAATGCGAGTTTGCAGAGGATTAACCT 4800  
Db ATAGGACGATAATAGGCTGGAAGCAATTAATGCGAGTTTGCAGAGGATTAACCT 4800  
Qy GTCTCAGGCTCTGATACAGAAACATCTGGTTCCTTCTATCATGTTTCTTCTCTAACA 4860  
Db GTCTCAGGCTCTGATACAGAAACATCTGGTTCCTTCTATCATGTTTCTTCTCTAACA 4860  
Qy TAAATCAGAGATACACCATGCTTCTGCTATAGATCTGTTTCATGTTTGGTGGTG 4920  
Db TAAATCAGAGATACACCATGCTTCTGCTATAGATCTGTTTCATGTTTGGTGGTG 4920  
Qy CTGTGTTAGCATTCCTATCTTGTATGGGATGACCTGTTGATCTGAGGCTTCTTTCAG 4980  
Db CTGTGTTAGCATTCCTATCTTGTATGGGATGACCTGTTGATCTGAGGCTTCTTTCAG 4980  
Qy TTAGTCTTCTTAACACCATTTATCTTCTTCCATTTGATCACCATGCGACACATGCTTC 5040  
Db TTAGTCTTCTTAACACCATTTATCTTCTTCCATTTGATCACCATGCGACACATGCTTC 5040  
Qy AGATACCTACTACAGTAGACACAGGCTTACCCCTGCTCAGGTTTCAGAAACAGTGAAG 5100  
Db AGATACCTACTACAGTAGACACAGGCTTACCCCTGCTCAGGTTTCAGAAACAGTGAAG 5100  
Qy AGGCTCATTTCCGCACTCTTCTTCTTTCAGAAATTTCTCAATATACAGTGGCTCCATTTG 5160  
Db AGGCTCATTTCCGCACTCTTCTTCTTTCAGAAATTTCTCAATATACAGTGGCTCCATTTG 5160  
Qy GGTGTGATTAATCTGGCTGATTTTGTGGGTCTCACTGAAGAAATGGCAATCCCTTATC 5220  
Db GGTGTGATTAATCTGGCTGATTTTGTGGGTCTCACTGAAGAAATGGCAATCCCTTATC 5220  
Qy TTGCTGTGCTGCAATTTTTCACATTTTCTTGGGTAACCTCGGCTGAGCACTGC 5280  
Db TTGCTGTGCTGCAATTTTTCACATTTTCTTGGGTAACCTCGGCTGAGCACTGC 5280  
Qy ATACCAGATCTGCGCTTACTAATCTGTTGAGCAAAACACCGTGGTCAAGTACCTTA 5340  
Db ATACCAGATCTGCGCTTACTAATCTGTTGAGCAAAACACCGTGGTCAAGTACCTTA 5340  
Qy ATTTGTTCTCTCTCTTCCAGGAATATTGGGATCTGTAAGGCCCTTGTCCAGAGGTGGT 5400  
Db ATTTGTTCTCTCTCTTCCAGGAATATTGGGATCTGTAAGGCCCTTGTCCAGAGGTGGT 5400  
Qy GTGCAGATCTGCGCTTACTAATCTGTTGAGCAAAACACCGTGGTCAAGTACCTTA 5460  
Db GTGCAGATCTGCGCTTACTAATCTGTTGAGCAAAACACCGTGGTCAAGTACCTTA 5460  
Qy GAAAAGAAATAGTTTGTATAGAGCTTCTGATGACTATAGTCTGCTCTGATCAAGCTT 5520  
Db GAAAAGAAATAGTTTGTATAGAGCTTCTGATGACTATAGTCTGCTCTGATCAAGCTT 5520  
Qy CTCAATTCAGGTGCCACGGTCTGAGATGATGAGCAAGCATCTGCTCTGCTGCTTT 5580  
Db CTCAATTCAGGTGCCACGGTCTGAGATGATGAGCAAGCATCTGCTCTGCTGCTTT 5580  
Qy TCTGTGGGGCTACTATGTTCTCAGAACATTTGCTGCGAGAAATTTGAAACGGGAG 5640  
Db TCTGTGGGGCTACTATGTTCTCAGAACATTTGCTGCGAGAAATTTGAAACGGGAG 5640  
Qy AGGTTGAGCTTGCATTTTTCAGCACTTCTACTGTGAGCGCGAGTCTGCAATTTCTTAA 5700  
Db AGGTTGAGCTTGCATTTTTCAGCACTTCTACTGTGAGCGCGAGTCTGCAATTTCTTAA 5700  
Qy AAATCAGAGAAATGCCAGTGGTCTCTGGTTGAAGGTAAAGCCAGAGGCTGTGCTATCAG 5760  
Db AAATCAGAGAAATGCCAGTGGTCTCTGGTTGAAGGTAAAGCCAGAGGCTGTGCTATCAG 5760

Qy 5761 CTCCTTACTTGGATGAATATGGAGAAACAGACCCCTGGCCCTGAAGAGGGGCAACCCCTTC 5820  
Db 5761 CTCCTTACTTGGATGAATATGGAGAAACAGACCCCTGGCCCTGAAGAGGGGCAACCCCTTC 5820  
Qy 5821 ATTTATCTCGTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACACACTGCAATTATAG 5880  
Db 5821 ATTTATCTCGTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACACACTGCAATTATAG 5880  
Qy 5881 AAGGATTTCTAGGAGCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTAC 5940  
Db 5881 AAGGATTTCTAGGAGCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTAC 5940  
Qy 5941 TGTGAGCTCCAACTCTCGCTCAAGCAATCAAAATGACGACACTAGTAAAGGCTGATTC 6000  
Db 5941 TGTGAGCTCCAACTCTCGCTCAAGCAATCAAAATGACGACACTAGTAAAGGCTGATTC 6000  
Qy 6001 AAAATTATGAAACATTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTCTTTGCTCCATGTC 6060  
Db 6001 AAAATTATGAAACATTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTCTTTGCTCCATGTC 6060  
Qy 6061 CAGGTTCACTTACATCAATTAATAATTTCTTAATGGAGTATTTGCTTCAATTAGCAACA 6120  
Db 6061 CAGGTTCACTTACATCAATTAATAATTTCTTAATGGAGTATTTGCTTCAATTAGCAACA 6120  
Qy 6121 TATGCTTTCACAGGAAAGGACATAGATCAATCTGTTTATGCTAGTAGTATTTCCAGGA 6180  
Db 6121 TATGCTTTCACAGGAAAGGACATAGATCAATCTGTTTATGCTAGTAGTATTTCCAGGA 6180  
Qy 6181 ATTTATTTCCCTTTCATTAATTTGCTCTCAATTTTATTTTATTCATCCACTTGGTAGAGAAG 6240  
Db 6181 ATTTATTTCCCTTTCATTAATTTGCTCTCAATTTTATTTTATTTTATTTTATTTTATTTT 6240  
Qy 6241 TCAGTCAAAACAGTTGTAGACATTTATGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300  
Db 6241 TCAGTCAAAACAGTTGTAGACATTTATGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300  
Qy 6301 GGTGTTTT 6308  
Db 6301 GGTGTTTT 6308

RESULT 2  
AV061886 5250 bp mRNA linear PRI 30-DEC-2002  
LOCUS Homo sapiens ubiquitin ligase E3 alpha-I mRNA, complete cds.  
DEFINITION  
ACCESSION AY061886  
VERSION AY061886.1 GI:27434483  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 5250)  
AUTHORS Han, H.Q. and Kwak, K.  
TITLE Full-length human ubiquitin ligase E3 alpha-I (E3 alpha)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 5250)  
AUTHORS Han, H.Q. and Kwak, K.  
TITLE Direct Submission  
JOURNAL Submitted (05-NOV-2001) Research, Amgen Inc., One Amgen Center  
Drive, Thousand Oaks, CA 91320, USA

FEATURES  
source  
1..5250  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..5250  
/codon\_start=1  
/product="ubiquitin ligase E3 alpha-I"  
/protein\_id="AAU32103.1"  
/db\_xref="GI:27434484"  
/translation="MADEEAGGTERMEISABELPQTPORLASWWDQQVDFYTFALHILA

OLVPRITPAEMDPDLKOEBSVQMSIPTPLEWYLFGEEDPDICLKLKHSAGAPOLCGRV  
FRSGRTTYSRCDCAIDPTCVLCMDCFQBSVHKHRYKQHTSTGCGFCDCGTEAWKTG  
PFCVNHGPRAGCTIKENSRNNEEVIQVARKIPSPVYKVVEMTIBEEKEPLPELQ  
IRBKRYCYVLFNDEHSHYHYVYSLOALDCELAQLHTTADKEGRRAVAGAY  
AACOKABEDIKSHSNVSOHPLHVEVLHSEIMAHOKFALRLGSMWNKMSYSDPQROI  
FCOACLRBPDSNPCLISRLMLDQAKYKARKILHLEIPSPFMEBYKXCLPAMFB  
VKYKLOLOKEYISDDHDSRISITALSVMQFTVPTLARIHERQNVISVITETLLEVL  
EVLDRNKNFNGOGYSQDKLGRVAVICDLKYLLISKPTIWERLBMQPLEGPRSPKI  
LTCMOGBEIRKQVQGHLEVDPDWEAATQMLKNLILMPOBWCACDEBELLVAYKE  
CHKAMRSTSPISSEKTVQSGHSLKTSYRVEDLVSHLPSRLTAGLHVLSR  
LGAVALHLEFMEQFQVEVLVEPLRCLVAQVVAEMWRNGLSL3QVFFYQDVK  
CREEMDYKIDIMLOICASLMDPNKLLVQRYELAEAFNKITSKDOODIKOYNVTL  
BEMLOVLYIVGERVVPVGVNVTKEBWTREIHLICLLEPMHPSAIAKXLPENRNKET  
GLEVINKVAFTKPGVSGHGYELKDSLQDPNNYFVHYSKTQHSKABHQKGRKQ  
ENKDALPPPPPEPCFAPFSKVINLNCIDIMYIILRTVFERAIDTDSNLMTGMLQMA  
PHILALQMLDEKQLOKAPSEEVTFYHKASRLGSSAMNTOMLLEKLGIPQLESGOK  
DMITWILQFDTVTKRLSKSLIVATTSKESIKNDEITHDKERKERRKAEARLHR  
OKIMQMSALQNFTEYTHKLMYNDTSEMPGKEDSIMEESTPAVSDYSRIALGPKRG  
SVTEKEVLTCILCOBEQEVKINNAMVLISACVOKSTALTOHREKPIELSGEALDPLM  
DPDLAYGTYTSGCHVMNAVCHOKYFPAVQLSSQORIHVDLPDLESGBYLPCLCKSLC  
NTVPIIPLOPKINSENADALQULLTLARMTQTLARIISGNIRHAKGNPPIPFN  
QGMGSTLEFHSILSPGVESSIKYSNISKEMVILFATTIYRIGLVKHPDPRVPM  
TSTCAFTQATLENLGEGLFGALQNRHNGKALMQPAVQORITCPOVLIOKHL  
VLLSVLPNLSKEDTPCLSLIDLPHVLGVAVLAPSLVDDPDVLPQSSVSSYNHL  
YLPHLITWAMLOILLVDVTCGLPLAQVEDSEAHSSAPFABISQYTSIGSDIPG  
WYLVWSLKNGITPYLRCAALPHYLGLVTPPELHNSABGSIYLSALSYLSTPLNLF  
LFQBYWDTVLPQORWADPALLNCLKQRTVYPRKRNLSLIELPDDYSCLLNQASH  
FRCPASADRRKHPVLCIFCGAILCSNICQOBIVNGEVEGACTPHALHCGAGVCIPL  
KIRCVVLVBGKARGCAVPAPLYDEYGETDPLKRGNPLHLRSRERYKHLHLVMOQHC  
IIEETARSOETNQMLPGFNWOLL"

ORIGIN

Query Match		83.2%;	Score 5250;	DB 9;	Length 5250;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 5250;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	696	ATGGCGGACGAGGAGCTGGAGTACTGAGAGATGGAATCAGCGCGGAGTTACCCAG	755		
DB	1	ATGGCGGACGAGGAGCTGGAGTACTGAGAGATGGAATCAGCGCGGAGTTACCCAG	60		
QY	756	ACCCCTCAGCGTCTGGCATCTGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG	815		
DB	61	ACCCCTCAGCGTCTGGCATCTGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG	120		
QY	816	CATCATTTGGCAAAATGGTCCAGAAATTTACTTTGCTGAAATGGACCAGACTTGGAA	875		
DB	121	CATCATTTGGCAAAATGGTCCAGAAATTTACTTTGCTGAAATGGACCAGACTTGGAA	180		
QY	876	AAGCAGGAGGAAGGTACAAATGTCAATATTCACTCCACTGGAATGCTATTATGGA	935		
DB	181	AAGCAGGAGGAAGGTACAAATGTCAATATTCACTCCACTGGAATGCTATTATGGA	240		
QY	936	GAGATCCAGATATTGTCTTAGAGAAATGAGACAGTGGAGCATTTTCAGCTTTGTGG	995		
DB	241	GAGATCCAGATATTGTCTTAGAGAAATGAGACAGTGGAGCATTTTCAGCTTTGTGG	300		
QY	996	AGGGTTTTCAAAGTGGAGAGCAACCTATTCTTCAGGGATGTGCAATGATCCACA	1055		
DB	301	AGGGTTTTCAAAGTGGAGAGCAACCTATTCTTCAGGGATGTGCAATGATCCACA	360		
QY	1056	TGTGTACTCTGTATGAGCTGCTTCAGGACAGTGTTCATAAAATCATCGTTACAAGATG	1115		
DB	361	TGTGTACTCTGTATGAGCTGCTTCAGGACAGTGTTCATAAAATCATCGTTACAAGATG	420		
QY	1116	CATACCTTCTACGGAGGAGGTTCTGTGACTGTGGAGACACAGAGGCATGGAAACTGGC	1175		
DB	421	CATACCTTCTACGGAGGAGGTTCTGTGACTGTGGAGACACAGAGGCATGGAAACTGGC	480		
QY	1176	CCTTTTGTGTAAATCATGAACCTCGAAGAGCAGGTACTATATAAAGAGAAATTCAGCGTGT	1235		
DB	481	CCTTTTGTGTAAATCATGAACCTCGAAGAGCAGGTACTATATAAAGAGAAATTCAGCGTGT	540		
QY	1236	CCGTTGAATGAGAGGTAAATGTCCAGCCAGGAAAAATATTTCCTTCAGTGAATAAATAT	1295		

DB	541	CCGTTGAATGAGAGGTAAATGTCCAGCCAGAAAAATATTTCCTTCAGTGAATAAATAT	600		
QY	1296	GTCTGAGAAATGACTATATGGGAAGAGAAAAAGAACTCCCTCTCGAACTCCAGATAAGG	1355		
DB	601	GTCTGAGAAATGACTATATGGGAAGAGAAAAAGAACTCCCTCTCGAACTCCAGATAAGG	660		
QY	1356	GAGAAAAATGAAGATACATATTGTCTCTTTTCAATGATGAACACCATTCATATGACCAAC	1415		
DB	661	GAGAAAAATGAAGATACATATTGTCTCTTTTCAATGATGAACACCATTCATATGACCAAC	720		
QY	1416	GTCTATACAGCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTGCATACC	1475		
DB	721	GTCTATACAGCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTGCATACC	780		
QY	1476	ACTGCCATTGACAAAGAGGGTCTGCGGGCTGTGTTAAAGCGGGAGCTTATGCTGCTTGCACG	1535		
DB	781	ACTGCCATTGACAAAGAGGGTCTGCGGGCTGTGTTAAAGCGGGAGCTTATGCTGCTTGCACG	840		
QY	1536	GAGCAAAAGGAGATATTAAGAGTCAATCAGAAAAATGTCTCTCAACATCCACTTCATGTA	1595		
DB	841	GAGCAAAAGGAGATATTAAGAGTCAATCAGAAAAATGTCTCTCAACATCCACTTCATGTA	900		
QY	1596	GAGTATTACACTCAGAGATTATGCTCATCAGAAAAATTTGCTTTCGCTCTTGGTCTCTGG	1655		
DB	901	GAGTATTACACTCAGAGATTATGCTCATCAGAAAAATTTGCTTTCGCTCTTGGTCTCTGG	960		
QY	1656	ATGAACAAAAATTTATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGGCAAGCATGCCCT	1715		
DB	961	ATGAACAAAAATTTATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGGCAAGCATGCCCT	1020		
QY	1716	AGAGAGAAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCA	1775		
DB	1021	AGAGAGAAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCA	1080		
QY	1776	AAGCTTTATAAGGTCGCCGTAAGATCCTTCATGAAATTTGCTTCAGCAGATTTTTTTATG	1835		
DB	1081	AAGCTTTATAAGGTCGCCGTAAGATCCTTCATGAAATTTGCTTCAGCAGATTTTTTTATG	1140		
QY	1836	GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAGATTTATAAAACAACTGCAG	1895		
DB	1141	GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAGATTTATAAAACAACTGCAG	1200		
QY	1896	AAAGATATATCAGTGATGATCATGACAGAGATATCTCTATAAATGCTCACTTCAGTTGAG	1955		
DB	1201	AAAGATATATCAGTGATGATCATGACAGAGATATCTCTATAAATGCTCACTTCAGTTGAG	1260		
QY	1956	ATGTTTACTGTTCTTACTCTGCTCGACATCTTATGGAAGCAGAAATGTTATCTCTGTC	2015		
DB	1261	ATGTTTACTGTTCTTACTCTGCTCGACATCTTATGGAAGCAGAAATGTTATCTCTGTC	1320		
QY	2016	ATTACTGAAACTCTGCTAGAAAGTTTACCTGAGTACTTGGCAGGAAACAATAAATTTCAAC	2075		
DB	1321	ATTACTGAAACTCTGCTAGAAAGTTTACCTGAGTACTTGGCAGGAAACAATAAATTTCAAC	1380		
QY	2076	TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAATATATGCACTTAAG	2135		
DB	1381	TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAATATATGCACTTAAG	1440		
QY	2136	TATATCTGTATCAGCAAAACCAATATGGAAGAGAAATTAAGNATGAGTTCTCTTGA	2195		
DB	1441	TATATCTGTATCAGCAAAACCAATATGGAAGAGAAATTAAGNATGAGTTCTCTTGA	1500		
QY	2196	GGTTTTGCATCTTTTTTGAAGATTTCTTACCTGTATGAGGGAATGGAAGAAATCCGAAGA	2255		
DB	1501	GGTTTTGCATCTTTTTTGAAGATTTCTTACCTGTATGAGGGAATGGAAGAAATCCGAAGA	1560		
QY	2256	CAGGTGGGCAACATTTGAAAGTGGATCTCTGATTTGGGAGGCTGCCATTTGCTATACAGATG	2315		
DB	1561	CAGGTGGGCAACATTTGAAAGTGGATCTCTGATTTGGGAGGCTGCCATTTGCTATACAGATG	1620		
QY	2316	CAATTTGAAGATATTTTACTCATGTTTCCAGAGTGGTGTCTTGTGATGAAGAACTCTTA	2375		

Db 1621 CAATTTGAAGAAATATTTTACTCATGTTTCCAGAGTGGTGTGCTGTGATGAAGAACTCTTA 1680  
Qy 2376 CTTGTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGGTGAGTACCAGTTTCATATCT 2435  
Db 1681 CTTGTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGGTGAGTACCAGTTTCATATCT 1740  
Qy 2436 AGTAGCAGACAGTAGTAGTAACAATCGTGTGCACATAGTTTGGAAAAAACAAGTCCCTACAGAGTA 2495  
Db 1741 AGTAGCAGACAGTAGTAGTAACAATCGTGTGCACATAGTTTGGAAAAAACAAGTCCCTACAGAGTA 1800  
Qy 2496 TCTGAGGATCTTGTGAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTCATGTA 2555  
Db 1801 TCTGAGGATCTTGTGAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTCATGTA 1860  
Qy 2556 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACATGCAATGAATTTGTGTCCTTTTGGAGACTTTT 2615  
Db 1861 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACTGCAATGAATTTGTGTCCTTTTGGAGACTTTT 1920  
Qy 2616 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTTGCTCTGGTGTGGTTGCCAGGTGCTT 2675  
Db 1921 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTTGCTCTGGTGTGGTTGCCAGGTGCTT 1980  
Qy 2676 GCTGAGATGTGGCGAAGAAATGGACTGTCTTTATTTAGCCAGGTGTTTATTTACCAAGAT 2735  
Db 1981 GCTGAGATGTGGCGAAGAAATGGACTGTCTTTATTTAGCCAGGTGTTTATTTACCAAGAT 2040  
Qy 2736 GTTAAGTGCAGAGAAAGAAATGTATGATAAAGATATCATCTTCAGATTTGGTGCACTT 2795  
Db 2041 GTTAAGTGCAGAGAAAGAAATGTATGATAAAGATATCATCATCTTCAGATTTGGTGCACTT 2100  
Qy 2796 TTAAATGGAATCCCAATAAGTCTTGTGTACTGTGTACTTCAGAGGTATGAACCTTGGCCAGGCT 2855  
Db 2101 TTAAATGGAATCCCAATAAGTCTTGTGTACTGTGTACTTCAGAGGTATGAACCTTGGCCAGGCT 2160  
Qy 2856 TTAAACAGACCATATCTACAAAAGCAGGATTTGATTTAAACAATAATAAATACATAA 2915  
Db 2161 TTAAACAGACCATATCTACAAAAGCAGGATTTGATTTAAACAATAATAAATACATAA 2220  
Qy 2916 GAAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGTGAGCGTTATGTACCTGGAGTGGGA 2975  
Db 2221 GAAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGTGAGCGTTATGTACCTGGAGTGGGA 2280  
Qy 2976 AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTCTTGTGATTTGCAATGAACCC 3035  
Db 2281 AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTCTTGTGATTTGCAATGAACCC 2340  
Qy 3036 ATGCCACACAGTCCCAATTTGACCAAAATTTTACCTGAGAAATGAATAAATGAATGCGCTTA 3095  
Db 2341 ATGCCACACAGTCCCAATTTGACCAAAATTTTACCTGAGAAATGAATAAATGAATGCGCTTA 2400  
Qy 3096 GAGAATGTCATAAACAAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGAGTT 3155  
Db 2401 GAGAATGTCATAAACAAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGAGTT 2460  
Qy 3156 TATGAACCTAAAGATGAATCATCTGAAGACTTTCATATGTACTTTTATCATTTACTCCAAA 3215  
Db 2461 TATGAACCTAAAGATGAATCATCTGAAGACTTTCATATGTACTTTTATCATTTACTCCAAA 2520  
Qy 3216 ACCCAGCATAGCAGGCTGAAATATGACAGAGAAAGAGAAAAACAAGAAACAAAGAT 3275  
Db 2521 ACCCAGCATAGCAGGCTGAAATATGACAGAGAAAGAGAAAAACAAGAAACAAAGAT 2580  
Qy 3276 GAAGCAATTCGCGCCACACACAGCTCCTGAAATTTCTGCCCTGCTTTTCAGCAAGTGAATTAAC 3335  
Db 2581 GAAGCAATTCGCGCCACACACAGCTCCTGAAATTTCTGCCCTGCTTTTCAGCAAGTGAATTAAC 2640  
Qy 3336 CTTTCTCAATGTGATATCATGATGTACATTTCTCAGACCGTATTTTGGCGGCAATAGAC 3395  
Db 2641 CTTTCTCAATGTGATATCATGATGTACATTTCTCAGACCGTATTTTGGCGGCAATAGAC 2700  
Qy 3396 ACAGATTTCTAACTTGTGGACCGAAGGATGTCTCCAAATGGCTTTTTCATATCTGGCATTTG 3455  
Db 2701 ACAGATTTCTAACTTGTGGACCGAAGGATGTCTCCAAATGGCTTTTTCATATCTGGCATTTG 2760

Qy 3456 GTTTTACTAGAGAGAGCAACAGCTTCAAAAAAGCTCTCTGAAGAAAGATTAACATTTTGAC 3515  
Db 2761 GTTTTACTAGAGAGAGCAACAGCTTCAAAAAAGCTCTCTGAAGAAAGATTAACATTTTGAC 2820  
Qy 3516 TTTTTATCATAGGCTTCAAGATTGGGAAGTTCCAGCCATGAATATACAAATGCTTTTGGAA 3575  
Db 2821 TTTTTATCATAGGCTTCAAGATTGGGAAGTTCCAGCCATGAATATACAAATGCTTTTGGAA 2880  
Qy 3576 AAACCTCAAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATAACGTGATACTTTCAG 3635  
Db 2881 AAACCTCAAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATAACGTGATACTTTCAG 2940  
Qy 3636 ATGTTTTCACACAGTGAAGCGAATTAAGAGAAAAATCTTGTGTTTAAATTTGTAGCAACACATCA 3695  
Db 2941 ATGTTTTCACACAGTGAAGCGAATTAAGAGAAAAATCTTGTGTTTAAATTTGTAGCAACACATCA 3000  
Qy 3696 GGATCGGAATCTATTAAGAATGATGAGATTAATCTCATGATAAAGAAAAAGCAGAAACGAAAA 3755  
Db 3001 GGATCGGAATCTATTAAGAATGATGAGATTAATCTCATGATAAAGAAAAAGCAGAAACGAAAA 3060  
Qy 3756 AGAAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGATCATGGCTCAGATGTCTGCCCTTA 3815  
Db 3061 AGAAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGATCATGGCTCAGATGTCTGCCCTTA 3120  
Qy 3816 CAGAAAAAATCTTATGAAAACTCATAAACTCATGTATGACAAATACATCAGAAATGCTCTGGG 3875  
Db 3121 CAGAAAAAATCTTATGAAAACTCATAAACTCATGTATGACAAATACATCAGAAATGCTCTGGG 3180  
Qy 3876 AAAGAGATTTCCATTTATGGAGAAAGAGACACCCAGCAGTCAGTGACTACTCTAGAAAT 3935  
Db 3181 AAAGAGATTTCCATTTATGGAGAAAGAGACACCCAGCAGTCAGTGACTACTCTAGAAAT 3240  
Qy 3936 GCTTTGGGTCTTAAACGGGTCTCATCTGTTACTGAAAAGAGAGGTGCTGACGTGCATCTT 3995  
Db 3241 GCTTTGGGTCTTAAACGGGTCTCATCTGTTACTGAAAAGAGAGGTGCTGACGTGCATCTT 3300  
Qy 3996 TGCCAGAGAAACAGGAGGTGAAAATAGAAATAATGCAATGTTATTCGGCTGTCTC 4055  
Db 3301 TGCCAGAGAAACAGGAGGTGAAAATAGAAATAATGCAATGTTATTCGGCTGTCTC 3360  
Qy 4056 CAGAAATCTACTGCTTTAAACCCAGCACAGGGGAAAAACCCATAGAACTCTCAGGAGAACCC 4115  
Db 3361 CAGAAATCTACTGCTTTAAACCCAGCACAGGGGAAAAACCCATAGAACTCTCAGGAGAACCC 3420  
Qy 4116 CTAGACCCACTTTTTCATGGATCCAGACTTGGCATATGGAACCTTATACAGGAGCTGTGCT 4175  
Db 3421 CTAGACCCACTTTTTCATGGATCCAGACTTGGCATATGGAACCTTATACAGGAGCTGTGCT 3480  
Qy 4176 CATGTAATGACAGCAGTGTCTGTCAGAAATATTTTGAAGCTGTACAGCTCAGCTCTCAG 4235  
Db 3481 CATGTAATGACAGCAGTGTCTGTCAGAAATGTTTGGAGCTGTACAGCTGAGCTCTCAG 3540  
Qy 4236 CAGCGCATTCATGTTGACCTTTTGGACTTGGAAAGTGGAGAAATATCTTTGGCCCTCTTTGC 4295  
Db 3541 CAGCGCATTCATGTTGACCTTTTGGACTTGGAAAGTGGAGAAATATCTTTGGCCCTCTTTGC 3600  
Qy 4296 AAATCTCTGTGCAATATCTGATCCCAATTTATTTCTTTGCAACCTCAAAAGATTAACAGT 4355  
Db 3601 AAATCTCTGTGCAATATCTGATCCCAATTTATTTCTTTGCAACCTCAAAAGATTAACAGT 3660  
Qy 4356 GAGAATGCAGATGCTTTTGTCTCAACTTTTGAACCTTGGCAGCGTGTGATACAGACTGTTCTG 4415  
Db 3661 GAGAATGCAGATGCTTTTGTCTCAACTTTTGAACCTTGGCAGCGTGTGATACAGACTGTTCTG 3720  
Qy 4416 GCCAGAAATATCAGGTTTAAATATAGACATGCTTAAAGGAGAAAAACCCAAATTTCTTATTTTC 4475  
Db 3721 GCCAGAAATATCAGGTTTAAATATAGACATGCTTAAAGGAGAAAAACCCAAATTTCTTATTTTC 3780  
Qy 4476 TTTTAAATCAGGAATGGGAGATTTCTACTTTGAGGTTCCATTTCCATCTCAGTTTGGCGTT 4535  
Db 3781 TTTTAAATCAGGAATGGGAGATTTCTACTTTGAGGTTCCATTTCCATCTCAGTTTGGCGTT 3840

4536 GAGTCTTCGATTAATAATTTCAATAGCATCAAGGAAATGGTTATTTCTCTTTCCCAACA 4595  
|||||  
3841 GAGTCTTCGATTAATAATTTCAATAGCATCAAGGAAATGGTTATTTCTCTTTCCCAACA 3900  
|||||  
4596 ATTATAGAAATGGATTGAAGTGCACCTGATGAAGGGATCCTCGAGTCCCCATGCTG 4655  
|||||  
3901 ATTATAGAAATGGATTGAAGTGCACCTGATGAAGGGATCCTCGAGTCCCCATGCTG 3960  
|||||  
4656 ACTTGAGACCTGCGCTTTCACTATCCAGGCAATTTGAAATCTATTTGGAGATGAAGGA 4715  
|||||  
3961 ACTTGAGACCTGCGCTTTCACTATCCAGGCAATTTGAAATCTATTTGGAGATGAAGGA 4020  
|||||  
4716 AAACCTCTCTTTGGAGCACTTCAAAATAGGAGCAATAATGCTGAAAGCAATTAATGCAG 4775  
|||||  
4021 AAACCTCTCTTTGGAGCACTTCAAAATAGGAGCAATAATGCTGAAAGCAATTAATGCAG 4080  
|||||  
4776 TTTGCAAGTTGCACAGAGGATTAACCTGCTCTCAGGCTCTGATACAGAAACATCTGGTTCTG 4835  
|||||  
4081 TTTGCAAGTTGCACAGAGGATTAACCTGCTCTCAGGCTCTGATACAGAAACATCTGGTTCTG 4140  
|||||  
4836 CTTCTATCAGTTGTTCTTCTTAACATAAATAACAGAAATACACCATGCTTCTGCTATA 4895  
|||||  
4141 CTTCTATCAGTTGTTCTTCTTAACATAAATAACAGAAATACACCATGCTTCTGCTATA 4200  
|||||  
4896 GATCTGTTTCATGTTTGGTGGGCTGTGTAGCATTCCTCATCTTGTATTGGATGAC 4955  
|||||  
4201 GATCTGTTTCATGTTTGGTGGGCTGTGTAGCATTCCTCATCTTGTATTGGATGAC 4260  
|||||  
4956 CTTGTTGATCTGCAGGCTTCTTTCAGTTAGTTCTTCTCTATAACCACTTTTATCTTCCAT 5015  
|||||  
4261 CTTGTTGATCTGCAGGCTTCTTTCAGTTAGTTCTTCTCTATAACCACTTTTATCTTCCAT 4320  
|||||  
5016 TTTGATCAACATGGCACATGTTTCTCAGATCTACTTACAGTAGACACAGGCTTACCCCTT 5075  
|||||  
4321 TTTGATCAACATGGCACATGTTTCTCAGATCTACTTACAGTAGACACAGGCTTACCCCTT 4380  
|||||  
5076 GCTCAGGTTCAAGAGACAGTAGAGGCTCTATTCCGATCTTCTTTCTTTGAGAAAT 5135  
|||||  
4381 GCTCAGGTTCAAGAGACAGTAGAGGCTCTATTCCGATCTTCTTTCTTTGAGAAAT 4440  
|||||  
5136 TCTCAATATACAAAGTGGCTCCATTGGTGTGTATTTCTGCTGTGTTTGGTCTCA 5195  
|||||  
4441 TCTCAATATACAAAGTGGCTCCATTGGTGTGTATTTCTGCTGTGTTTGGTCTCA 4500  
|||||  
5196 CTGAAGAATGGCATCACCCCTTATCTTCGCTGTGCTGATGTTTTCACATATTACTT 5255  
|||||  
4501 CTGAAGAATGGCATCACCCCTTATCTTCGCTGTGCTGATGTTTTCACATATTACTT 4560  
|||||  
5256 GGGGTAACTCCGCTGAGGAACCTGCATACCAATTTCTGCAAGAGGAGTACAGTGCACTC 5315  
|||||  
4561 GGGGTAACTCCGCTGAGGAACCTGCATACCAATTTCTGCAAGAGGAGTACAGTGCACTC 4620  
|||||  
5316 TGTAGCTATCTATCTTTACCTACAAATTTGCTGCTCTCCAGGAATTTGGGATCT 5375  
|||||  
4621 TGTAGCTATCTATCTTTACCTACAAATTTGCTGCTCTCCAGGAATTTGGGATCT 4680  
|||||  
5376 GTAAGGCCCTTGTCTCCAGAGTGGTGTGAGATCTTCCCTTACTAACTGTTTGAAGCAA 5435  
|||||  
4681 GTAAGGCCCTTGTCTCCAGAGTGGTGTGAGATCTTCCCTTACTAACTGTTTGAAGCAA 4740  
|||||  
5436 AAAAACCCGCTGGTACCGCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGAC 5495  
|||||  
4741 AAAAACCCGCTGGTACCGCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGAC 4800  
|||||  
5496 TATAGTGTCTCTGATCAAGCTTCTCAATTTTCAGTGCCACGCTCTGAGATGAG 5555  
|||||  
4801 TATAGTGTCTCTGATCAAGCTTCTCAATTTTCAGTGCCACGCTCTGAGATGAG 4860  
|||||  
5556 CGAAAGCATCTGTCTCTGCTCTTCTGTGGGCTTACTATGTTCTCAGAACATTTGC 5615  
|||||  
4861 CGAAAGCATCTGTCTCTGCTCTTCTGTGGGCTTACTATGTTCTCAGAACATTTGC 4920  
|||||  
5616 TCCAGGAAATTTGTGAAACGGGGAAGAGGTTGGAGCTTGCAATTTTTCACGCACTTCACTGT 5675  
|||||

4921 TCCAGGAAATTTGTGAACGGGAGAGGTTGGAGCTTGCAATTTTTCACCACTTCACTGT 4980  
|||||  
5676 GGAGCCGGAGTCTGCATTTTCTCTAAAAATCAGAGAATCCGAGTGGTCTCTGGTTGAAGT 5735  
|||||  
4981 GGAGCCGGAGTCTGCATTTTCTCTAAAAATCAGAGAATCCGAGTGGTCTCTGGTTGAAGT 5040  
|||||  
5736 AAAGCCAGAGGCTGTGCTTATCCAGTCTCTTACTTGGATGAATATGGAGAAACAGACCT 5795  
|||||  
5041 AAAGCCAGAGGCTGTGCTTATCCAGTCTCTTACTTGGATGAATATGGAGAAACAGACCT 5100  
|||||  
5796 GGCCTGAAGAGGCGCAACCCCTTCAATTTATCTCTGAGCGGTATCGGAAGCTTCAATTTG 5855  
|||||  
5101 GGCCTGAAGAGGCGCAACCCCTTCAATTTATCTCTGAGCGGTATCGGAAGCTTCAATTTG 5160  
|||||  
5856 GTCTGGCAACAACTGCAATTAAGAAGATTTGCTAGGAGCCAGAGACTAATCAGATG 5915  
|||||  
5161 GTCTGGCAACAACTGCAATTAAGAAGATTTGCTAGGAGCCAGAGACTAATCAGATG 5220  
|||||  
5916 TTATTTGGATTCACTGGCAGTTACTGTGA 5945  
|||||  
5221 TTATTTGGATTCACTGGCAGTTACTGTGA 5250  
|||||  
RESULT 3  
AR487565 5205 bp DNA linear PAT 14-MAY-2004  
LOCUS Sequence 18 from patent US 6706505.  
DEFINITION AR487565  
ACCESSION AR487565  
VERSION AR487565.1 GI:47252815  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 5205)  
AUTHORS Han, H.-Q. and Kwak, K.  
TITLE Human E3.alpha. ubiquitin ligase family  
JOURNAL Patent: US 6706505-A 18 16-MAR-2004;  
FEATURES Location/Qualifiers  
1..5205  
source /organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 81.4%; Score 5136; DB 6; Length 5205;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;  
QY 696 ATGCGGACGAGGAGGCTGAGGTTACTGAGAGATGGAATCAGCGCGGAGTTACCCAG 755  
DB 1 ATGCGGACGAGGAGGCTGAGGTTACTGAGAGATGGAATCAGCGCGGAGTTACCCAG 60  
QY 756 ACCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 815  
DB 61 ACCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 120  
QY 816 CATCATTTGGCAAAATTTGGTGCAGAAATTTACTTTGTGTAATGGACCCAGACTTGGAA 875  
DB 121 CATCATTTGGCAAAATTTGGTGCAGAAATTTACTTTGTGTAATGGACCCAGACTTGGAA 180  
QY 876 AAGCAGGAGGAAAGTGTACAAATTTCAATTTCACTCCACTCGAATGGTACTTATTGGA 935  
DB 181 AAGCAGGAGGAAAGTGTACAAATTTCAATTTCACTCCACTCGAATGGTACTTATTGGA 240  
QY 936 GAGATCCAGATTTTGTCTTGAAGAAATTTGAGACAGTGGAGCATTTTCAAGCTTTGTGG 995  
DB 241 GAGATCCAGATTTTGTCTTGAAGAAATTTGAGACAGTGGAGCATTTTCAAGCTTTGTGG 300  
QY 996 AGGGTTTTCAAAAGTGGAGAGCAACCTTATTTCTCAGGGATTTGTCAATTTGATCCAA 1055  
DB 301 AGGGTTTTCAAAAGTGGAGAGCAACCTTATTTCTTGCAGGGATTTGTCAATTTGATCCAA 360  
QY 1056 TGTGTAATTTGTGAAACGGGGAAGAGGTTGGAGCTTGCAATTTTTCACGCACTTCACTGT 1115  
|||||



|||||  
Db 361 TGTGTACTCTGTATGCACTGCTTCCAGGACAGTGTTCATATAAATCATCGTTACAGATG 420  
QY 1116 CATACCTTCTACGGAGGAGGTTCTGTGTGTGTGAGACACAGAGGCGATGGAAAACCTGGC 1175  
Db 421 CATACCTTCTACGGAGGAGGTTCTGTGTGTGTGTGAGACACAGAGGCGATGGAAAACCTGGC 480  
QY 1176 CCTTTTGTGTAAATCATGAACCTCGGAAGAGGAGTACTATATAAGAGAAATTCACGCTGT 1235  
Db 481 CCTTTTGTGTAAATCATGAACCTCGGAAGAGGAGTACTATATAAGAGAAATTCACGCTGT 540  
QY 1236 CGTTTGAATGAAGAGGTAATTTGTCACAGCCAGGAAAATATTTCTCTCAGTGTATAAAATAT 1295  
Db 541 CGTTTGAATGAAGAGGTAATTTGTCACAGCCAGGAAAATATTTCTCTCAGTGTATAAAATAT 600  
QY 1296 GTCTAGAAAATGACTATATGGGAAGAGGAAAAGAACTGCTCTCGAACTCCAGATAAGG 1355  
Db 601 GTCTAGAAAATGACTATATGGGAAGAGGAAAAGAACTGCTCTCGAACTCCAGATAAGG 660  
QY 1356 GAGAAAATGAAGATACCTATTTGTGTCTTTTCAATGATGAACACCATTCATATGACCCAC 1415  
: : : : :  
Db 661 KNR-----YYCNDHSHSYDH 675  
QY 1416 GTCATATACAGCTTACAAAGAGCTCTTGACTGTGTGAGCTCGCAGAGGCCAGTTGCATACC 1475  
Db 676 GTCATATACAGCTTACAAAGAGCTCTTGACTGTGTGAGCTCGCAGAGGCCAGTTGCATACC 735  
QY 1476 ACTGCCATTGACAAAGAGGCTCGTGGGCTGTATAAGCGGGAGCTTATGTCTGTGGCAG 1535  
Db 736 ACTGCCATTGACAAAGAGGCTCGTGGGCTGTATAAGCGGGAGCTTATGTCTGTGGCAG 795  
QY 1536 GAAGCAAGGAAGATATAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACCTTCATGTA 1595  
Db 796 GAAGCAAGGAAGATATAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACCTTCATGTA 855  
QY 1596 GAAGTATTACACTCAGAGATATGCTCATACAGAAAATTTGCTTGGCTCTGGTTCTCTGG 1655  
Db 856 GAAGTATTACACTCAGAGATATGCTCATACAGAAAATTTGCTTGGCTCTGGTTCTCTGG 915  
QY 1656 ATGAAACAAATTTATGAGCTATTCAAGTGACTTTTAGGCAGATCTTTTGGCAAGCATGCTTT 1715  
Db 916 ATGAAACAAATTTATGAGCTATTCAAGTGACTTTTAGGCAGATCTTTTGGCAAGCATGCTTT 975  
QY 1716 AGAGAAGAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTTAAATGCTTTGGGATGCA 1775  
Db 976 AGAGAAGAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTTAAATGCTTTGGGATGCA 1035  
QY 1776 AAGCTTTTATAAGGTGCCGTAAAGATCCTTCATGAATGATCTTCAGCAGATTTTTTTATG 1835  
Db 1036 AAGCTTTTATAAGGTGCCGTAAAGATCCTTCATGAATGATCTTCAGCAGATTTTTTTATG 1095  
QY 1836 GAGATGGAAATACAAAAAATCTTTTGTCTATGGAAATTTGTGAAGTATATAAACAATGCA 1895  
Db 1096 GAGATGGAAATACAAAAAATCTTTTGTCTATGGAAATTTGTGAAGTATATAAACAATGCA 1155  
QY 1896 AAAGAATATATCATGTATGATCATGACAGAGTATCTCTATAACCTGCACCTTCAGTTGAG 1955  
Db 1156 AAAGAATATATCATGTATGATCATGACAGAGTATCTCTATAACCTGCACCTTCAGTTGAG 1215  
QY 1956 ATGTTTACTGTTTCTTACTCTGCTCGACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTC 2015  
Db 1216 ATGTTTACTGTTTCTTACTCTGCTCGACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTC 1275  
QY 2016 ATTACTGAAATCTGCTAGAACTTTTACCTGAGTACTTTGGACAGGAAACAATAAATCAAC 2075  
Db 1276 ATTACTGAAATCTGCTAGAACTTTTACCTGAGTACTTTGGACAGGAAACAATAAATCAAC 1335  
QY 2076 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTATATGTCACCTTAAG 2135  
Db 1336 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTATATGTCACCTTAAG 1395  
QY 2136 TATATCTCTGATCAGCAAAACCCCAATATATGACAGAAAGATTAAGAAATGCAAGTTCCTTGAA 2195  
|||||

Db 1396 TATATCTCTGATCAGCAAAACCCCAATATGACAGAAAGATTAAGAAATGCAAGTTCCTTGAA 1455  
QY 2196 GGTTTTCGATCTTTTGTGAAGATTTCTTACCTGTATGAGGAAATGGAGAATCCGAGA 2255  
Db 1456 GGTTTTCGATCTTTTGTGAAGATTTCTTACCTGTATGAGGAAATGGAGAATCCGAGA 1515  
QY 2256 CAGGTGGGCAACACATTTGAAGTGGATCCTCATTTGGGAGGCTGCCATTTGCTTATACAGATG 2315  
Db 1516 CAGGTGGGCAACACATTTGAAGTGGATCCTCATTTGGGAGGCTGCCATTTGCTTATACAGATG 1575  
QY 2316 CAATGAAGAAATATTTTATCTCATGTTCCAGAGTGGTGTGTGTGTGTATGAAGAACTCTTA 2375  
Db 1576 CAATGAAGAAATATTTTACTCATGTTCCAGAGTGGTGTGTGTGTATGAAGAACTCTTA 1635  
QY 2376 CTTGTGGCTTATAAGAAATGTCACAAAGCTGTGTATGAGGTGAGTACAGTTTTCATATCT 2435  
Db 1636 CTTGTGGCTTATAAGAAATGTCACAAAGCTGTGTATGAGGTGAGTACAGTTTTCATATCT 1695  
QY 2436 AGTAGCAAGACAGTAGTACAAATCGTGTGGACATAGTTTGGAAAACAAAGTCCCTACAGAGTA 2495  
Db 1696 AGTAGCAAGACAGTAGTACAAATCGTGTGGACATAGTTTGGAAAACAAAGTCCCTACAGAGTA 1755  
QY 2496 TCTGAGGATCTTGTAAAGCATACTGTGCCACTCTCTAGGACCCCTTGTCTTCAATGTA 2555  
Db 1756 TCTGAGGATCTTGTAAAGCATACTGTGCCACTCTCTAGGACCCCTTGTCTTCAATGTA 1815  
QY 2556 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACTGCATGAATTTGTCTTTTGGAGACTTTT 2615  
Db 1816 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACTGCATGAATTTGTCTTTTGGAGACTTTT 1875  
QY 2616 CAAGTAGAGGTACTAGTGGAAATATCCTTTAGCTGTCTGTGTGTGTGTGTGTGTGTGT 2675  
Db 1876 CAAGTAGAGGTACTAGTGGAAATATCCTTTAGCTGTCTGTGTGTGTGTGTGTGTGTGT 1935  
QY 2676 GCTGAGATGTGGCGAAGAAATGGAATGCTCTTATTTAGCCAGGTGTGTGTATTTACCAAGAT 2735  
Db 1936 GCTGAGATGTGGCGAAGAAATGGAATGCTCTTATTTAGCCAGGTGTGTGTATTTACCAAGAT 1995  
QY 2736 GTTAAGTGCAGAGAAATGTATGATATAAGATATCATCATGCTTCAGATTTGGTGCATCT 2795  
Db 1996 GTTAAGTGCAGAGAAATGTATGATATAAGATATCATCATGCTTCAGATTTGGTGCATCT 2055  
QY 2796 TTAATGGAATCCCAATTAAGTCTTGTGTACTGTGTACTTCAGAGGTATGAATTTGCCAGGCT 2855  
Db 2056 TTAATGGAATCCCAATTAAGTCTTGTGTACTGTGTACTTCAGAGGTATGAATTTGCCAGGCT 2115  
QY 2856 TTTAAACAGACCATATCTACAAAAGACAGGATTTGATTAAACAATATAATACATAATA 2915  
Db 2116 TTTAAACAGACCATATCTACAAAAGACAGGATTTGATTAAACAATATAATACATAATA 2175  
QY 2916 GAAGAAATGCTTCAGTCTCTCATCTATATTGTGGGTGAGCGTTATGTACCTGGAGTGGGA 2975  
Db 2176 GAAGAAATGCTTCAGTCTCTCATCTATATTGTGGGTGAGCGTTATGTACCTGGAGTGGGA 2235  
QY 2976 AATGTGACCAAGAGAGGTCACAATGAGAGAAATCATTCACCTTGTCTTTGCAATGAAACCC 3035  
Db 2236 AATGTGACCAAGAGAGGTCACAATGAGAGAAATCATTCACCTTGTCTTTGCAATGAAACCC 2295  
QY 3036 ATGCCACACAGTGCCTATTTGCCAAAAATTTTACTGAGAATGAAAATAATGAAACTGGCTTTA 3095  
Db 2296 ATGCCACACAGTGCCTATTTGCCAAAAATTTTACTGAGAATGAAAATAATGAAACTGGCTTTA 2355  
QY 3096 GAGATGTCTATAAACAAAGTGGCCACATTTTAAGAAACAGGTGTATCAGGCCATGAGTT 3155  
Db 2356 GAGATGTCTATAAACAAAGTGGCCACATTTTAAGAAACAGGTGTATCAGGCCATGAGTT 2415  
QY 3156 TATGAATCTAAAAGAGTGAATCACTGAAAGACTTCAATATGTACTTTTATCTTACTCCAAA 3215  
Db 2416 TATGAATCTAAAAGAGTGAATCACTGAAAGACTTCAATATGTACTTTTATCTTACTCCAAA 2475  
QY 3216 ACCCAGCATACAGGCTGAAACATATGCAAGAGAAAGAGAGAAACAAAGAAACAAAGAT 3275  
Db 2476 ACCCAGCATACAGGCTGAAACATATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2535  
|||||

QY 3276 GAAGCAATTCGCCGCCACCACTCTCTGAAATTTCTGCCCTGCTTTTCAGGCAAGTGATTAAC 3335  
DB 2536 GAAGCAATTCGCCGCCACCACTCTCTGAAATTTCTGCCCTGCTTTTCAGGCAAGTGATTAAC 2595  
QY 3336 CTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGCAATAGAC 3395  
DB 2596 CTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGCAATAGAC 2655  
QY 3396 ACAGATTTCTAACTGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATTTCTGGCAATG 3455  
DB 2656 ACAGATTTCTAACTGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATTTCTGGCAATG 2715  
QY 3456 GGTTTACTAGAGAGAGCAACAGCTTCAAAGAGCTCTGAAGAGTCTGAAGAGAGTAACTTTGAC 3515  
DB 2716 GGTTTACTAGAGAGAGCAACAGCTTCAAAGAGCTCTGAAGAGTCTGAAGAGAGTAACTTTGAC 2775  
QY 3516 TTTTATCATAGGCTTCAAGATTTGGAGTTGAGAGTTCCAGCCATGATATACAAATGCTTTTGGAA 3575  
DB 2776 TTTTATCATAGGCTTCAAGATTTGGAGTTGAGAGTTCCAGCCATGATATACAAATGCTTTTGGAA 2835  
QY 3576 AAACCTCAAAGGAATTTCCCGATTTAGAGGCCAGAGAGCATGATTAACCTGGATATCTTCAG 3635  
DB 2836 AAACCTCAAAGGAATTTCCCGATTTAGAGGCCAGAGAGCATGATTAACCTGGATATCTTCAG 2895  
QY 3636 ATGTTTGAACAGTGAAGCGATTAAGAGAAAATCTTGTTTAAATTTGTAGCAACCATCA 3695  
DB 2896 ATGTTTGAACAGTGAAGCGATTAAGAGAAAATCTTGTTTAAATTTGTAGCAACCATCA 2955  
QY 3696 GGATCGGAATCTTAATTAAGAGTATGAGATTAATCTCATGATTAAGAGAAAAGCAGAGCAAAA 3755  
DB 2956 GGATCGGAATCTTAATTAAGAGTATGAGATTAATCTCATGATTAAGAGAAAAGCAGAGCAAAA 3015  
QY 3756 AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3815  
DB 3016 AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3075  
QY 3816 CAGAAAACTTCATTTGAAACTCATAACTCATGATATGACATACATCAGAAAATGCTGGG 3875  
DB 3076 CAGAAAACTTCATTTGAAACTCATAACTCATGATATGACATACATCAGAAAATGCTGGG 3135  
QY 3876 AAGAGAGATTCATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3935  
DB 3136 AAGAGAGATTCATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3195  
QY 3936 GCTTTGGGTCTTAACGGGTCCATCTGTACTGAAAAGAGAGTCTGACGTGATCTT 3995  
DB 3196 GCTTTGGGTCTTAACGGGTCCATCTGTACTGAAAAGAGAGTCTGACGTGATCTT 3255  
QY 3996 TGCCAAAGAGAACAGAGAGTGAATAAGAAAATATGCCATGGTATTAATCGGCTGTGTC 4055  
DB 3256 TGCCAAAGAGAACAGAGAGTGAATAAGAAAATATGCCATGGTATTAATCGGCTGTGTC 3315  
QY 4056 CAGAAATCTATGCTCCCTTAACCCAGCAGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 4115  
DB 3316 CAGAAATCTATGCTCCCTTAACCCAGCAGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 3375  
QY 4116 CTAGACCCACTTTTTCATGATCCAGCTTGGCATATGAACTTATACAGGAGAGCTGTGCT 4175  
DB 3376 CTAGACCCACTTTTTCATGATCCAGCTTGGCATATGAACTTATACAGGAGAGCTGTGCT 3435  
QY 4176 CATGTAATGACGAGAGTGTCTGGCAGAGATTTTGAAGCTGTACAGCTGAGCTCTCAG 4235  
DB 3436 CATGTAATGACGAGAGTGTCTGGCAGAGATTTTGAAGCTGTACAGCTGAGCTCTCAG 3495  
QY 4236 CAGGCAATTCATGTTGACTTTTGAATTTGAGAGTGAAGTATATCTTTGGCTCTTTG 4295  
DB 3496 CAGGCAATTCATGTTGACTTTTGAATTTGAGAGTGAAGTATATCTTTGGCTCTTTG 3555  
QY 4296 AAATCTGTGCAATCTGATCCCAATTTCTTTGCAACCTCCTCAAGAGTAAACAGT 4355  
DB 3556 AAATCTGTGCAATCTGATCCCAATTTCTTTGCAACCTCCTCAAGAGTAAACAGT 3615

QY 4356 GAGATGAGATGCTCTTGTCTCAACTTTTGACCCCTGGCAGCGTGATACAGACTGTCTG 4415  
DB 3616 GAGATGAGATGCTCTTGTCTCAACTTTTGACCCCTGGCAGCGTGATACAGACTGTCTG 3675  
QY 4416 GCCAGAAATACAGGTTATAATATAGACATGCTAAAGGAGAAAACCCAAATTCCTATTTTC 4475  
DB 3676 GCCAGAAATACAGGTTATAATATAGACATGCTAAAGGAGAAAACCCAAATTCCTATTTTC 3735  
QY 4476 TTTAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATTCATCTCTGAGTTTGGCGTT 4535  
DB 3736 TTTAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATTCATCTCTGAGTTTGGCGTT 3795  
QY 4536 GAGCTTCGATTAATATTTCAATAGCATCAAGGAATGGTATTTCTCTTTGCCCAACA 4595  
DB 3796 GAGCTTCGATTAATATTTCAATAGCATCAAGGAATGGTATTTCTCTTTGCCCAACA 3855  
QY 4596 ATTTATAGAAATGGAGTTGAAAGTGCCTCATGAAAGGGATCTCTCGAGTCCCAATGCTG 4655  
DB 3856 ATTTATAGAAATGGAGTTGAAAGTGCCTCATGAAAGGGATCTCTCGAGTCCCAATGCTG 3915  
QY 4656 ACCTGGAGCACCTGCGCTTTCACTATCCAGGCAATTTGAAAATCTATTGGGAGATGAAGA 4715  
DB 3916 ACCTGGAGCACCTGCGCTTTCACTATCCAGGCAATTTGAAAATCTATTGGGAGATGAAGA 3975  
QY 4716 AAACTCTGTTTGGAGCATCTTCAAAAATAGGCGAGCATTAATGGTCTGAAAAGCATTAATGCG 4775  
DB 3976 AAACTCTGTTTGGAGCATCTTCAAAAATAGGCGAGCATTAATGGTCTGAAAAGCATTAATGCG 4035  
QY 4776 TTTGAGTTGACAGAGAGATTTACCTCTCAGGTTCTGATACAGAAACATCTGCTTCTG 4835  
DB 4036 TTTGAGTTGACAGAGAGATTTACCTCTCAGGTTCTGATACAGAAACATCTGCTTCTG 4095  
QY 4836 CTCTCTCAGTTGTTCTTCTTAACATAAAATCAGAGATACACCAATGCTTCTGCTCTATA 4895  
DB 4096 CTCTCTCAGTTGTTCTTCTTAACATAAAATCAGAGATACACCAATGCTTCTGCTCTATA 4155  
QY 4896 GATCTGTTTCAATGTTTGGTGGTCTGTTAGCATTTCCCATCTCTGTATTTGGAGTAC 4955  
DB 4156 GATCTGTTTCAATGTTTGGTGGTCTGTTAGCATTTCCCATCTCTGTATTTGGAGTAC 4215  
QY 4956 CCTGTTGATCTGAGCCCTTCTTCAAGTATGTTCTTCTATACCACTTATCTCTTCCAT 5015  
DB 4216 CCTGTTGATCTGAGCCCTTCTTCAAGTATGTTCTTCTATACCACTTATCTCTTCCAT 4275  
QY 5016 TTGATCACCATGCGCACATGCTTTCAGATATCTACTTACAGTAGACACAGGCTTACCCCTT 5075  
DB 4276 TTGATCACCATGCGCACATGCTTTCAGATATCTACTTACAGTAGACACAGGCTTACCCCTT 4335  
QY 5076 GCTCAGGTTCAAGAGACAGTGAAGAGGCTCATTCGGCATCTTCTTTTTCGAGAAAT 5135  
DB 4336 GCTCAGGTTCAAGAGACAGTGAAGAGGCTCATTCGGCATCTTCTTTTTCGAGAAAT 4395  
QY 5136 TCTCAATATACAGTGGTCCATTTGGGTGATATCTCTGGCTGCTTCTTCTGGGTCTCA 5195  
DB 4396 TCTCAATATACAGTGGTCCATTTGGGTGATATCTCTGGCTGCTTCTTCTGGGTCTCA 4455  
QY 5196 CTGAAGATGTCATCACCCCTTATCTTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCT 5255  
DB 4456 CTGAAGATGTCATCACCCCTTATCTTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCT 4515  
QY 5256 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGAGAGAGAGTACAGTGCATC 5315  
DB 4516 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGAGAGAGAGTACAGTGCATC 4575  
QY 5316 TGTAGCTATCTATCTTTTACCTACAAATTTGTTTCTGCTTCTTCTTCTTCTTCTTCTTCT 5375  
DB 4576 TGTAGCTATCTATCTTTTACCTACAAATTTGTTTCTGCTTCTTCTTCTTCTTCTTCTTCT 4635  
QY 5376 GTAAAGGCTTGTCTCCAGAGGTGGTGTGAGATCTCTGCTTCTTCTTCTTCTTCTTCTTCT 5435  
DB 4636 GTAAAGGCTTGTCTCCAGAGGTGGTGTGAGATCTCTGCTTCTTCTTCTTCTTCTTCTTCT 4695  
QY 5436 AAAAACAACCGTGTGAGTACCCCTAGAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGAC 5495

Db	4696	AAAAACCGTGGTCAAGGTACCTAGAAAAAGAAATAGTTTGATAGAGCTTCTCGATGAC	4755
QY	5496	TATAGTGGTCTCCTGAATCAAGCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGAG	5555
Db	4756	TATAGTGGTCTCCTGAATCAAGCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGAG	4815
QY	5556	CGAAGACATCCTGCTCCTGCTTCTGCTGGGGCTACTATGTTCTCAGAACATTTGC	5615
Db	4816	CGAAGACATCCTGCTCCTGCTTCTGCTGGGGCTACTATGTTCTCAGAACATTTGC	4875
QY	5616	TGCCAGGAAATTTGTGAACGGGGAAGAGGTGTGAGCTTGCATTTTTCAGCACTTCACTGT	5675
Db	4876	TGCCAGGAAATTTGTGAACGGGGAAGAGGTGTGAGCTTGCATTTTTCAGCACTTCACTGT	4935
QY	5676	GGAGCGGAGTGTGCATTTTCTAAATACAGAGATGCCGAGTGGTCTCGTTGAAGGT	5735
Db	4936	GGAGCGGAGTGTGCATTTTCTAAATACAGAGATGCCGAGTGGTCTCGTTGAAGGT	4995
QY	5736	RAAGCCAGAGGTGTGCCTATCCAGCTCCTTACTTGGATGATATGGAGAACACACCT	5795
Db	4996	RAAGCCAGAGGTGTGCCTATCCAGCTCCTTACTTGGATGATATGGAGAACACACCT	5055
QY	5796	GGCCTGAAGAGGGGCAACCCCTTTCATTTATCTCGTGAGCGGTATCGGAAGCTCCATTTG	5855
Db	5056	GGCCTGAAGAGGGGCAACCCCTTTCATTTATCTCGTGAGCGGTATCGGAAGCTCCATTTG	5115
QY	5856	GTCTGGCAACAACATGCAATATAGAGATGTCTAGAGCCAGAGACTAATCAGATG	5915
Db	5116	GTCTGGCAACAACATGCAATATAGAGATGTCTAGAGCCAGAGACTAATCAGATG	5175
QY	5916	TTATTGGATTCACTGGCAGTTACTGTGA	5945
Db	5176	TTATTGGATTCACTGGCAGTTACTGTGA	5205

RESULT 4	
AF525401	
LOCUS	AF525401 Homo sapiens UBRI E3a ligase mRNA, partial cds. PRI 01-JAN-2003
DEFINITION	Homo sapiens UBRI E3a ligase mRNA, partial cds.
ACCESSION	AF525401
VERSION	AF525401.1 GI:27451603
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Dgany,O., Avidan,N., Delaunay,J., Krasnov,T., Shalmon,L., Shalev,H., Eidelitz-Markus,T., Kapelushnik,J., Cattán,D., Pariente,A., Tulliez,M., Cretien,A., Schischmanoff,P.O., Iolascon,A., Fibach,E., Koren,A., Rossler,J., Le Merrer,M., Yaniv,I., Zaizov,R., Ben-Asher,E., Olender,T., Lancet,D., Beckmann,J.S. and Tamary,H.
TITLE	Congenital dyserythropoietic anemia type I is caused by mutations in codanin-1
JOURNAL	Am. J. Hum. Genet. 71 (6), 1467-1474 (2002)
MEDLINE	22340442
PUBMED	12434312
REFERENCE	2 (bases 1 to 5128)
AUTHORS	Ben-Asher,E.
TITLE	Direct Submission
JOURNAL	Submitted (27-JUN-2002) Molecular Genetics, The Weizmann Institute of Science and the Schneider Medical Center, Hertzfel Street, Rehovot 76100, Israel
FEATURES	Location/Qualifiers
source	1..5128
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="15"
	/map="15q13"
	/cell_type="erythroid"

CDS

<1..>5128	
/codon_start=1	
/product="UBRI E3a ligase"	
/protein_id="AA014997.1"	
/db_xref="GI:27451604"	
/translation="LPQTQRLASWMDQVDFTYAPLHLHLAQLVPEIYFAEMDPLEK QESVQMSIPTPLEWYFGEDPDI CLEKLXSGAFOLCGRVFKSGETTYSCRDCAIDP TCVLCMDPQDSVHKHRYKMTSTGGFCDCGDTAWEKTPGFCVNHPEGRAGTKEN SRCPNREVLVQARKIEPVLKYKVENTIMEERKELPELQIREKNERYVCVLPNDEH LSYDHYVLSQALDCBLAQAQLHTTALIDEGRRAVKAGAYAAQCEAKEDIKSHENV SQPHLHVHSEIMAHQFALRGLSWMNKIMSYSDRFQFCQACLRDEPDSENPL IIRMLWDKLYGARKILHELIFSPFMEBYKCLFAMFVYTKYQLOKLEYISDDHD RSISTALSVMQFTVPTLRLHLEBEQVIVITETLEVLPEYLDKNNKFFQYSQD KLVGVAVICDLKYLISKPTIWTRELMOFLGFRSLKILTCWQMEERLRVVGOKH LEVDPWEAAIAIQMLKNILLMPOEWCACDELLAVAYKECHKAVMRCSPITSSSK TVVSCGHSLETYSYRVSEDLVLSHLPLSRTLAGLHVRLSRLGAVSRLHEFVSFEDFQ VLVVLYPLRCLVLVAQVAAEMRRNGLSLSISQVYYODVKREEMDDKDLIMLOIGA SLMDPNKFLLLVQLYELAEAFNPTISTKQDQLIKQYNTLIEMLQVLIYIVGSRYP SGVNTKEVTWREI IHLCTIEMPHSAIANKLPENNNETGLNNKINVPFKKPGV GNGVYELKDBSLKDFNNYFVYKSTQHSKAEHMKRKRQENDEALPPPEPCFP AFKVINLLNCDIMMYILRTVPERAIDTDSNLMTGMLQMAFHILALGLLEKQLOK APBEVTFDFTHKASRLGSSAMNIMLEKLGKIPQLEGQKTMITWILOMDFVRLR EKSCLIVATTSGSESINKDEITHDKERAKRKAARLHRKQIMAQMSALQKPIET HKLMYDNTSEMPGKEDSIMEESTPAVSDYSRIALPKRGPSTVEKEVLTILCOBEQ EVKLENNAMVLSACVQKSTALTQHRKPIELSGEALDPLFMDPLAYTGYTILSCGHM HAVCWQYFRAVOLSSOORIHVDLDESLEYLCLPKSLCNTVPIPIPLQPKINSF NADALAILLAWIOTVLARIISGYNIRHAKGNPIPIPFNQMGDSTLHFSILSGF VESSIKYSNISKENVILPATIIRIGILKVPDERDPRVPLMTNSTCAFTQAIENLIG DEGLPFLGQNRHNGLLKMPAQVITCPQVLIQKHLVRLLSVLPNKSIEDTP CLLSIDLFLVLVGAFLAPSLYDWDPLQPSVSSYNNHLYLPHLITMAHMLQILLT VDTGLPLAQVEDSEEAHSASFPABISQVTSIGICDIPGWLWYVSLKNGKITPYLRWC ADPLFVLLGVTPPELHTNSAEGBYSALCSYLSLXNLLPLLOBYWVDTVRLPLQRC ADPALNCLKQKTVVRYPRKENSILIEPDYSCLLNOASHFRCPRSADDERKHPVLC LFCGAILCSQNICQBEIVNGEEVGACIFPHALHCGAGVCIFLKIICREVRVLVEGKARGC AYPAYILDEYGETDPLKLRGNPLHFLVSGTEAPLVROOHC"	

ORIGIN

Query Match	80.6%;	Score	5085.6;	DB	9;	Length	5128;
Best Local Similarity	99.8%;	Pred. No.	0;				
Matches	5121;	Conservative	1;	Mismatches	6;	Indels	3;
Gaps	3;						
QY	747	TTATCCCCAGAGCCCTCAGCGTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATACT	806				
Db	1	TTATCCCCAGAGCCCTCAGCGTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATACT	60				
QY	807	GCTTCTTGTGATCATTTGGGCAAAATGCTGTCAGAAAATTTACTTCTGGAATGGACCCA	866				
Db	61	GCTTCTTGTGATCATTTGGGCAAAATGCTGTCAGAAAATTTACTTCTGGAATGGACCCA	120				
QY	867	GACTTGGAAAAAGCAGGAGGAAAGTGACAAATGTCATATTTCACTCCACTCGAATGGTAC	926				
Db	121	GACTTGGAAAAAGCAGGAGGAAAGTGACAAATGTCATATTTCACTCCACTCGAATGGTAC	180				
QY	927	TTATTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAAGCAGTGGAGCAATTCAG	986				
Db	181	TTATTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAAGCAGTGGAGCAATTCAG	240				
QY	987	CTTTGTGGAGGGTTTTCAAAAGTGGAGAGACAACCTATTTCTTGAGGGATTTGCAATT	1046				
Db	241	CTTTGTGGAGGGTTTTCAAAAGTGGAGAGACAACCTATTTCTTGAGGGATTTGCAATT	300				
QY	1047	GATCCACATCTGTACTCTGTATGACTGCTTCCAGGACAGTGTTCATATAAAATCATCGT	1106				
Db	301	GATCCACATCTGTACTCTGTATGACTGCTTCCAGGACAGTGTTCATATAAAATCATCGT	360				
QY	1107	TACAAGATGCATCTTCTACTGGAGGGGTTCTGTGACTGTGGAGACACAGAGCATGG	1166				
Db	361	TACAAGATGCATCTTCTACTGGAGGGGTTCTGTGACTGTGGAGACACAGAGCATGG	420				
QY	1167	AAAACCTGCCCTTTTGTGTAATCATGAACTGGAGAGCAGGTACTATATAAGAGAT	1226				
Db	421	AAAACCTGCCCTTTTGTGTAATCATGAACTGGAGAGCAGGTACTATATAAGAGAT	480				





Db 2641 GCATAGACACAGATTCTAACTTGTGACCGAAGGGATGCTCCAAATGGCTTTTTCATATT 2700  
Qy 3447 CTGGCAATTGGGTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCCCTGGAAGAACTA 3506  
Db 2701 CTGGCAATTGGGTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCCCTGGAAGAACTA 2760  
Qy 3507 ACATTTGACTTTTATCATAAAGGCTTCAAGATTGGGAAGTTTCAAGCCATGAATATACAAATG 3566  
Db 2761 ACATTTGACTTTTATCATAAAGGCTTCAAGATTGGGAAGTTTCAAGCCATGAATATACAAATG 2820  
Qy 3567 CTTTTGGAAAAAATCTCAAAAGGAATCCCAAGTTAGAAAGCCAGAAAGCAATGATATACGTTG 3626  
Db 2821 CTTTTGGAAAAAATCTCAAAAGGAATCCCAAGTTAGAAAGCCAGAAAGCAATGATATACGTTG 2880  
Qy 3627 ATACTTCAGATGTTTTGACACAGTGAAGCGATTAAAGAGAAATCTTGTTTAAATTTGTAGCA 3686  
Db 2881 ATACTTCAGATGTTTTGACACAGTGAAGCGATTAAAGAGAAATCTTGTTTAAATTTGTAGCA 2940  
Qy 3687 ACCACATCAGGATCGGAATCTATTAAAGAAATGATGAGATTACTCATGATATAAGAAAAAGCA 3746  
Db 2941 ACCACATCAGGATCGGAATCTATTAAAGAAATGATGAGATTACTCATGATATAAGAAAAAGCA 3000  
Qy 3747 GAAAGAAAAAGAAAGCTGAGCTAGGCTACATGCGCCAGAAAGATCATGGCTCAGATG 3806  
Db 3001 GAAAGAAAAAGAAAGCTGAGCTAGGCTACATGCGCCAGAAAGATCATGGCTCAGATG 3060  
Qy 3807 TCTGCTTTACAGAAAAAATCTTCAATGAACTCATAACTCATGATATGACAAATACATCAGAA 3866  
Db 3061 TCTGCTTTACAGAAAAAATCTTCAATGAACTCATAACTCATGATATGACAAATACATCAGAA 3120  
Qy 3867 ATGCTCGGAAAAAGAAATCTTCAATGAGGAAGAGAGCACCCAGCAGTCACTGACTAC 3926  
Db 3121 ATGCTCGGAAAAAGAAATCTTCAATGAGGAAGAGAGCACCCAGCAGTCACTGACTAC 3180  
Qy 3927 TCTAGAAATGCTTTGGGTCCTAAACGGGGTCCATCTGTTACTGAAAAAGAGGTGCTGACG 3986  
Db 3181 TCTAGAAATGCTTTGGGTCCTAAACGGGGTCCATCTGTTACTGAAAAAGAGGTGCTGACG 3240  
Qy 3987 TGCACTCTTGGCAAGAAAGACAGAGGTGAATAATAGAAATAATGCCATGGTATTATCG 4046  
Db 3241 TGCACTCTTGGCAAGAAAGACAGAGGTGAATAATAGAAATAATGCCATGGTATTATCG 3300  
Qy 4047 GCCTGTGTCCAGAAATCTACTGCCCTTAAACAGCAGCAGGGAAAAACCCATAGAACTCTCA 4106  
Db 3301 GCCTGTGTCCAGAAATCTACTGCCCTTAAACAGCAGCAGGGAAAAACCCATAGAACTCTCA 3360  
Qy 4107 GGAGAGCCCTAGACCCACTTTTCATGGATCCAGACTTGGCATATGGAATCTTATACAGA 4166  
Db 3361 GGAGAGCCCTAGACCCACTTTTCATGGATCCAGACTTGGCATATGGAATCTTATACAGA 3420  
Qy 4167 AGCTGTGTGATGTAATGACCGCAGTGTGCTGGCAGAGTATTTTGAAGCTGTACAGCTG 4226  
Db 3421 AGCTGTGTGATGTAATGACCGCAGTGTGCTGGCAGAGTATTTTGAAGCTGTACAGCTG 3480  
Qy 4227 AGCTGTGACGCGCATTCATGTTGACCTTTTGTGCTTGAAGTGGAGAAATCTTTGTC 4286  
Db 3481 AGCTGTGACGCGCATTCATGTTGACCTTTTGTGCTTGAAGTGGAGAAATCTTTGTC 3540  
Qy 4287 CCTCTTTGCAAACTCTGTGTCAATATCTGTGATCCCACTTTTCTTTTGAACCTCAAAAG 4346  
Db 3541 CCTCTTTGCAAACTCTGTGTCAATATCTGTGATCCCACTTTTCTTTTGAACCTCAAAAG 3600  
Qy 4347 ATAAACAGTGAGAAATGCAATGCTCTTGTCTGAACTTTTGAACCTGGCAGGTTGATACAG 4406  
Db 3601 ATAAACAGTGAGAAATGCAATGCTCTTGTCTGAACTTTTGAACCTGGCAGGTTGATACAG 3660  
Qy 4407 ACTGTTTGGCCAGAAATATCAGGTTATATAATAAGACATGCTAAAGAGGAAAAACCAATT 4466  
Db 3661 ACTGTTTGGCCAGAAATATCAGGTTATATAATAAGACATGCTAAAGAGGAAAAACCAATT 3720  
Qy 4467 CCTATTCTTTTAAATCAAGGAATGGAGATCTTACTTTGGAGTTCCATTTCCATCTCGAGT 4526  
Db 3721 CCTATTCTTTTAAATCAAGGAATGGAGATCTTACTTTGGAGTTCCATTTCCATCTCGAGT 3780

Qy 4527 TTTGGCGTTGAGTCTTCGATTAAATATTCAAAATAGCATCAAGGAAATGGTTATTCTCTTT 4586  
Db 3781 TTTGGCGTTGAGTCTTCGATTAAATATTCAAAATAGCATCAAGGAAATGGTTATTCTCTTT 3840  
Qy 4587 GCCACAAACAATTTATAGAAATTTGGATTGGAATGCACTGATGAAAGGATCTTCGAGTC 4646  
Db 3841 GCCACAAACAATTTATAGAAATTTGGATTGGAATGCACTGATGAAAGGATCTTCGAGTC 3900  
Qy 4647 CCATGCTGACCTGGAGCACCTGGCGCTTTCATCTATCCAGGCAATTTGAAAATCTATTGGGA 4706  
Db 3901 CCATGCTGACCTGGAGCACCTGGCGCTTTCATCTATCCAGGCAATTTGAAAATCTATTGGGA 3960  
Qy 4707 GATGAAGGAAAACTCTGTTTGGAGCACCTTCAAAATAGGAGCATTAATGGTCTCAAAAGCA 4766  
Db 3961 GATGAAGGAAAACTCTGTTTGGAGCACCTTCAAAATAGGAGCATTAATGGTCTGAAAGCA 4020  
Qy 4767 TTAATGCAAGTTTGCAGTTGCAACAGAGATTAACCTGTCTCAGGCTCGTGATACAGAAACAT 4826  
Db 4021 TTAATGCAAGTTTGCAGTTGCAACAGAGATTAACCTGTCTCAGGCTCGTGATACAGAAACAT 4080  
Qy 4827 CTGGTTGCTCTTCTATCAGTTGTTCTTCTAAACATTAATAACAGAGATACACCATGCCCTT 4886  
Db 4081 CTGGTTGCTCTTCTATCAGTTGTTCTTCTAAACATTAATAACAGAGATACACCATGCCCTT 4140  
Qy 4887 CTGTCTATAGATCTGTTTCAATGTTTGGTGGTCTGTGTTAGCATTCCTCATCTCTGTAT 4946  
Db 4141 CTGTCTATAGATCTGTTTCAATGTTTGGTGGTCTGTGTTAGCATTCCTCATCTCTGTAT 4200  
Qy 4947 TGGGATGACCTGTGTGATCTGCAGCCTTCTTCAAGTTAGTTCTTCTATAAACACCTTTAT 5006  
Db 4201 TGGGATGACCTGTGTGATCTGCAGCCTTCTTCAAGTTAGTTCTTCTATAAACACCTTTAT 4260  
Qy 5007 CTCTTCCATTTGATCACCATGGGACACATGCTTCAAGTACTACTTACAGTACACACAGC 5066  
Db 4261 CTCTTCCATTTGATCACCATGGGACACATGCTTCAAGTACTACTTACAGTACACACAGC 4320  
Qy 5067 CTACCCCTTGTCTCAGGTTCAAGAAAGACAGTGAAGAGGCTCATTCGGCATCTTCTTCTTT 5126  
Db 4321 CTACCCCTTGTCTCAGGTTCAAGAAAGACAGTGAAGAGGCTCATTCGGCATCTTCTTCTTT 4380  
Qy 5127 GCAGAAATTTCTCAATATACAAAGTGGCTCCATGGGTGTGATATCTTGGCTGGTATTG 5186  
Db 4381 GCAGAAATTTCTCAATATACAAAGTGGCTCCATGGGTGTGATATCTTGGCTGGTATTG 4440  
Qy 5187 TGGGTCTCAGTGAAGATGGCATCACCCCTTATCTTGGCTGTGATCTTCTTTTCCAC 5246  
Db 4441 TGGGTCTCAGTGAAGATGGCATCACCCCTTATCTTGGCTGTGATCTTCTTTTCCAC 4500  
Qy 5247 TATTTACTTGGGGTAACTCCGCCCTGAGGAACCTGCATACCAATTTGCAAGAAAGGAGTAC 5306  
Db 4501 TATTTACTTGGGGTAACTCCGCCCTGAGGAACCTGCATACCAATTTGCAAGAAAGGAGTAC 4560  
Qy 5307 AGTGACCTCTGATGCTATCTATCTTTTACCTAACAATTTGTTCTGCTCTTCCAGGAATAT 5366  
Db 4561 AGTGACCTCTGATGCTATCTATCTTTTACCTAACAATTTGTTCTGCTCTTCCAGGAATAT 4620  
Qy 5367 TGGGATACTGTAAAGCCCTTGTCTCAGAGGTGGTGTGAGATCTGCGCTTACTATAACTGT 5426  
Db 4621 TGGGATACTGTAAAGCCCTTGTCTCAGAGGTGGTGTGAGATCTGCGCTTACTATAACTGT 4680  
Qy 5427 TTGAAGCAAAAAAACCACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTT 5486  
Db 4681 TTGAAGCAAAAAAACCACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTT 4740  
Qy 5487 CCTGATGCTATAGCTGCTCCTGAAATCAAGCTTCTCAATTTCAAGTGGCCCGCTGCA 5546  
Db 4741 CCTGATGCTATAGCTGCTCCTGAAATCAAGCTTCTCAATTTCAAGTGGCCCGCTGCA 4800  
Qy 5547 GATGATGAGCAAAAGCATCTGCTCTGCGCTTTTCTGCGGCTTACTACTATGTTCTCAG 5606  
Db 4801 GATGATGAGCAAAAGCATCTGCTCTGCGCTTTTCTGCGGCTTACTACTATGTTCTCAG 4860

QY 5607 AACATTTGCTGCCAGGAATTTGTAACGGGGAAGGTTGGAGCTTGCAATTTTTCACGCA 5666  
DB 4861 AACATTTGCTGCCAGGAATTTGTAACGGGGAAGGTTGGAGCTTGCAATTTTTCACGCA 4920  
QY 5667 CTTCACTGTGGAGCCGGAGTCTGCAATTTTCCATAAATCAGAGAAATCCGAGTGGTCTG 5726  
DB 4921 CTTCACTGTGGAGCCGGAGTCTGCAATTTTCCATAAATCAGAGAAATCCGAGTGGTCTG 4980  
QY 5727 GTTGAAGGTAAGCCAGAGGCTGTGCTATCCAGTCCCTTACTTGGATGAATATCGAGAA 5786  
DB 4981 GTTGAAGGTAAGCCAGAGGCTGTGCTATCCAGTCCCTTACTTGGATGAATATCGAGAA 5040  
QY 5787 ACAGACCCCTGGCCTGAGAGGGGCAACCCCTTCATTTATCTGTCGAGGCTATCGGAAG 5846  
DB 5041 ACAGACCCCTGGCCTGAGAGGGGCAACCCCTTCATTTTCTCGTGAAGCGGTA-CGGAG 5098  
QY 5847 CTCCTATTGGTCTGGCAACCAACATGCAATTA 5877  
DB 5099 CTCCTATTGGTCTGGCAACCAACATGCAATTA 5128

## RESULT 5

AR030784  
LOCUS AR030784 6395 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5861312.  
ACCESSION AR030784  
VERSION AR030784.1 GI:5943998  
KEYWORDS

## SOURCE

Unknown.

## ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 6395)

AUTHORS Varshavsky, A. and Kwon, Y. Tae.

TITLE Nucleic acid encoding mammalian UBR1

JOURNAL Patent: US 5861312-A 1 19-JAN-1999;

FEATURES

source

1..6395

/organism="unknown"

/mol\_type="unassigned DNA"

## ORIGIN

Query Match 70.3%; Score 4435.4; DB 6; Length 6395;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;

6;

QY 590 TTCAGGGCGCGTCGTAAGGTGTCCTGCTGTC-CGCGACGGCCACAGGTTCCGCT 648

DB 8 TTCAGGGCGCGTCGTAAGGTGTCCTGCTGTCGCGTCGGGCGGCCACAGGTTCCGCT 67

QY 649 TGCCTCTGCGCGGGGTGCGCAACTGCGAGCGTCAAGTTTCCCTCAAGATGGCGGACGAGG 708

DB 68 AGCTGGCGCGGGGTGCGCAACTGCGGGGCTTCGTTTCCCTTAAGATGGCGGACGAGG 127

QY 709 AGGCTGGAGGTACTGAGAGGATGGAATACAGCGGGGATTCACCCAGACCCCTCAGGTC 768

DB 128 AGATGAGCGCGCGGAGAGGATGGAAGTCAAGTTCAGCGCGGAGCGCTCCCTGGCGCGGACGCGC 187

QY 769 TGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTCTGTCATCATTTGGCAC 828

DB 188 CGGCATCTGTTGGGATCAGCAAGTTGATTTTATCTGCTTCTTATCATTTTGGCAC 247

QY 829 AATTGGTCCAGAAATTTACTTTTGTGTAATGGACCCAGACTTGGAAAGCAGGAGGAAA 888

DB 248 AATTAGTCCAGAAATTTATTTTGCTGAGATGGACCCAGATTTGGAAAGCAGGAGAGA 307

QY 889 GTGTACAAATGTCAATATTCATCTCACTGGAATGTAATTTTGGAGAAATCCAGATA 948

DB 308 GTGTACAGATGTCAATATTCATCTCTCTTGGAGTGGTACTTATTTGGAGAGGATCCGATA 367

QY 949 TTTCGTTAGAGAAATTTGAAGCAGTGGGCAATTTTCAGCTTTTGGGAGGCTTTCAAAA 1008

DB 368 TTTCGTTAGAGAAATTTAAAAACAGTGGAGGCTTCAGTTGTGTGGGAAAGGTTTCAAAA 427

QY 1009 GTGGAGAGACAACCTATTCTTGAGGGGATTTGTCGAATTTGATCCAAATGTTGTACTCTGTA 1068  
DB 428 GTGGAGAGACAACCTATTCTTGAGGGGATTTGTCGAATTTGATCCAAATGTTGTACTCTGTA 487  
QY 1069 TGGACTGCTCCAGGACAGTGTTCATAAAAATCATCGTTTACAAGATGATCTCTACTG 1128  
DB 488 TGGACTGCTTCAAAAGTATGTTTCATAAAAACCATCTGTTACAGATGATCTCTACTG 547  
QY 1129 GAGGAGGGTCTGTGTAATCTGTGAGAGACAGAGGCAATGGAATACTGGCCCTTTTGTGTA 1188  
DB 548 GAGGAGGGTCTGTGTAATCTGTGAGAGACAGAGGCAATGGAATACTGGCCCTTTTGTGTA 607  
QY 1189 ATCATGAACCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248  
DB 608 ATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 667  
QY 1249 AGGTAATTTGTCAGAGCCAGGAAATATTTCTTCAGTGTATAAATATGTCGTGAGAAATGA 1308  
DB 668 AGGTAATTTGTCAGAGCCAGGAAATATTTCTTCAGTGTATAAATATGTCGTGAGAAATGA 727  
QY 1309 CTATATGGGAAGAGAGAAAGAACTGCTCTGTAATCCAGATAGGGAGAGAAATGAAA 1368  
DB 728 CTATATGGGAAGAGAGAAAGAAAGGAAATGCTCTCTGTAATAGGCAGATAGGGAGAGAAATGAAC 787  
QY 1369 GATACATATGTCCTTTTCAATGATGAACACCAATTCATATGACACACATCATATACAGCC 1428  
DB 788 GATACATATGTCCTTTTCAACGATGAGCACCATTGCTATGATCATGTCATACAGTC 847  
QY 1429 TACAAGAGCTCTTGACTGTGAGCTCGCAGAGGGCCAGTTCATACCACTGCCATTGACA 1488  
DB 848 TGCAGAGAGCTCTAGATTGCGAGCTTGCAGAGGCAAGCTGCACAGCTGCCATCGACA 907  
QY 1489 AAGAGGGTCTGCGGGCTGTTAAAGCGGAGCTTATGCTGTTGCGAGAAAGCAAGAGAG 1548  
DB 908 AAGAGGGTCTGCGGGCTGTTCAAGCAGAGTGTATGCCACTTGCAGAGAGCAAGAGAGG 967  
QY 1549 ATATAAGAGTCAATTCAGAAATGTCCTCAACATGCCACTTCATGTAGTAGAGTATTACACT 1608  
DB 968 ATATAAGAGTCAATTCAGAGAACTCTCTCAGCACCCTCCATGTGGAAGTGTGCACT 1027  
QY 1609 CAGAGATTTATGGCTCATCAGAAATTTGCTTTGCGTCTTGCTTCTGAGTGAACAAAATTA 1668  
DB 1028 CCGTGGTTATGGCTCACCAGAAATTCGCTCTGCGCTTGGCTCTCGGATGAACAAAATTA 1087  
QY 1669 TGAGCTATTCAAGTGAATTTAGGAGAGTCTTTTGCAGAGCATGCTTATAGAGAAACCTG 1728  
DB 1088 TGAGCTATTCAAGTGAATTTAGACAGATATTTTGCAGGCTGCTCTGTAGAGAGACCTG 1147  
QY 1729 ACTCGAGAAATCCCTGCTCATAGAGAGGTTAAATGCTTTGGGATGCAAGCTTTATAAAG 1788  
DB 1148 GCTCTGAAATCCCTGCTTATAAGCAGACTAATGCTTTGGGATGCAAAACCTTTATAAAG 1207  
QY 1789 GTGCGGCTAGATCTTCATGAATTTGATCTTCAGCAGTCTTTTATGAGATGGAATACA 1848  
DB 1208 GTGCGGCTAGATCTTCATGAATTTGATCTTTAGTAGTCTTTTATGAGATGGAATACA 1267  
QY 1849 AAAAATCTTTTGTATGGAATTTTGTGAAGTATTAATAACACTGCGAGAAAGATATATCA 1908  
DB 1268 AAAAATCTTTTGTATGGAATTTTGTGAAGTATTAATAACACTGCGAGAAAGATATCA 1327  
QY 1909 GTGATGATCATGACAGAGATATCTTATACTGCACTTTTCAGTTCAGATGTTTACTGTTTC 1968  
DB 1328 GCGAGCAGCAGAGAGAGAGAGATCTCCATTAACCGCCCTGTCTCGTGCAGATGCTCACCGTCC 1387  
QY 1969 CTACTCTGGCTCGACATCTTATGAAGAGAGAGATGTTTCTCTGCTCATTTACTGAAACTC 2028  
DB 1388 CGACCTTTGGCCCGGATCTTTATGAAGAGAGAGAGATGTTTCTCTGCTCATTTACTGAAACTC 1447  
QY 2029 TGCTAGAGATTTTACCTGAGTACTTGGACAGAGAAATTAATTCAACTTCAGAGGCTTATA 2088  
DB 1448 TGCTAGAGATTTTACCTGAGTACTTGGACAGAGAGAGATTAATTCAACTTCAGAGGCTTATA 1507  
QY 2089 GCCAGGACAAATTTGGGAGAGATATATGTCAGTAAATATGTCAGCTAAAGATATATCTGATCA 2148



Qy	4300	CTCTGTGCAATAC	TGTGTGATCCCA	ATTAATCTCTTTGCA	AACTCTCAAAAGATA	AAACAAGTGAGA	4359
Db	3728	CTCTCTGCAA	CACTGTCTATCC	CTCTTTGACGCGC	GACAGAGATCA	CAACAGTGAGA	3787
Qy	4360	ATGCAGATGCT	CTTGTCTCAACT	TTTTTGACCCCTTG	GCACCGTGATAC	GACATGTTCTTGCCCA	4419
Db	3788	ATGCGGAGGCT	CTTGTCTCAACT	TTTTTGACCTTGG	CCCGTGATACAG	ACTGTCTCTGCCA	3847
Qy	4420	GAATATCAGG	TTATAATAAG	ACATGCTAAAGG	AGAAAACC---	CAATTCCTATTTTCT	4476
Db	3848	GAATATCGGG	TTATAATAAG	CAATGCTAAAGG	AGAAAGCC	CAGCAGTTCCTGTCTTGT	3907
Qy	4477	TTAATCAAGG	AAATGGGAGAT	TCTACTTTTGG	AGTTCCATTCAT	CCTGAGTTTTGGCGTTG	4536
Db	3908	TTAATCAAGG	AAATGGGGAT	TCAACTTTTGA	TTTTCATTCAT	CCTGAGTTTTGGAGTTC	3967
Qy	4537	AGTCTTCGAT	TAAATATTT	CAAAATAGCAT	CAAGGAAATGG	TATTCTCTTTTGCCACAACA	4596
Db	3968	AGTCTTCGG	TGAATATTT	CAAAATAGTAT	CAAGGAAATGG	TCTCTCTTTGCGCAACAACA	4027
Qy	4597	TTTTAAGAA	TTGGATGAA	AGTGCCA	CCTGATGAAGG	ATCTCTCGAGTCCCCATGCTGA	4656
Db	4028	TTTTACAGAA	TTGGCTGAAG	TGGCTCTCTGATGA	ACTAGACCC	CACGAGTGCCTCATGTA	4087
Qy	4657	CCTGGAGCA	CTGCGCTTCA	CTATCCAGGCA	ATTTGAATATCT	ATTGGGAGATGAAGGAA	4716
Db	4088	CCTGGAGCA	CGTGTGCGT	TCACTATCCAC	ATCAGGCAATCG	AAATCTATTGGGAGATGAAGGAA	4147
Qy	4717	AACCTCTGT	TTGGAGCACT	TCAAAATAGG	CAGCAATAATGG	TCTCAAAAGCATTTAATGCAGT	4776
Db	4148	AACCTCTAT	TTGGAGCACT	TCAAAATAG	CATAGCGTCTG	NAAGCGCTAATGCAGT	4207
Qy	4777	TTGCAGTTG	CAAGAGATTA	CTGTCTCTCAG	TCTGATACAGAA	CAATCTGTGTTCTG	4836
Db	4208	TTGCAGTTG	CAAGAGGCT	ACTCTGCCCT	CAGGCTCTGATAC	CAAAACATCTGGCTCGGC	4267
Qy	4837	TTCTATCAG	TGTGTTCTT	CCTAACATAA	ATCAGAGATAT	CACCATGCGTCTCTGCTATAG	4896
Db	4268	TCCTGT	CAGTTATTCT	TCTAACCTG	CAATCAGAAAT	ATACACGAGGCTCTCTGTCTGTG	4327
Qy	4897	ATCTGTTTC	ATGTTTGGT	GGTGCTGTGTT	TAGCATTTCC	CAATCCTGTGATTTGGGATGACC	4956
Db	4328	ATCTTCTCC	ATGTTCTGG	TGGCGCAGTCT	TAGGTTCCCA	CTCTTGATTTGGGATGACA	4387
Qy	4957	CTGTTGAT	CTGCAGCCT	TTCTCAGTTAG	TCTCTCTATA	ACCACCTTTATCTCTTCCATT	5016
Db	4388	CCGTGGAT	CTGCAGCGT	CGCACTTAGT	TCTCATATAAC	CACTCTATCTCTTCCATC	4447
Qy	5017	TGATACAC	ATNGGCACAT	CGTTCAGAT	ACTACTTAC	GATAGACA-----CAG	5064
Db	4448	TGATACCA	ATGGCGCA	ATGCTTCAGAT	ACTCTTCA	CAACAGATACAGATCTGTCTTCAG	4507
Qy	5065	GCCTTAC	CCCTTGCTC	CAGGTTCA	GAAGAAGCAGT	GAAAGGCTCATTCGCGCATCTCTTCTCT	5124
Db	4508	GGCCGCGCT	TGCTGAGGGT	GAAGAGATG	TGAGGAGGCT	GTGCTGTGCATCTGCTTCTCT	4567
Qy	5125	TTGCAGAAA	TTTCTCAAT	ATATACAA	GTGGTCTC	ATTTCTCTGGCTGTGTTAT	5184
Db	4568	TTTGTGAAG	TGTCGCAG	CAACACAG	CGGCTCTCA	TGGGTGCGGTGCTCCGCGCTGTGATCC	4627
Qy	5185	TGTGGGTCT	CACTGAAG	ATGGCATCAC	CCCTTATCTT	CGCTGTGCTGCATTTGTTTTCC	5244
Db	4628	TGTGGCTCT	CCCTGAGGA	AGGCAATAC	CCCTTACCT	TCCGCTGTGCTGTGCACTGCTTTTCC	4687
Qy	5245	ACTATTAT	TTCTGGGT	TAACCTCG	CTCAGGAAC	TGCAATACAAATTTCTGCAGAAAGGAGT	5304
Db	4688	ACTATTACT	TTGGAGT	AGCTCG	CTCAGAA	CTGTTTGGCAATTTCTGCTGAAGGAAAT	4747
Qy	5305	ACAGTGCA	CTCTGTAG	CTATCTAT	CTTTA	CCTACAAATTTGTCTGCTCTTCCAGAAAT	5364
Db	4748	TCAGTGCA	CTCTGTAG	CTATCTAT	CTTTACCA	CAAAATTTGTCTGCTCTTTCAGAAAT	4807

Qy	5365	ATTGGGATAC	TGTAAGGCCCTT	TGCTCAGAGGTGGTGTGCAGATCCTTGCCCTTACTAAACT	5424
Db	4808	ATTGGGATAC	CAATAAGGCCCTT	TACTACAGAGGTGGTGTGCAGATCCTTGCCCTTACTCAAGT	4867
Qy	5425	GTTTGAAGCA	AAAAAACACCGTGGT	CAGGTACCTAGAAAAAGAAATAGTTTTCATAGAGC	5484
Db	4868	CTTTGAAGCA	AAAAAGTGTGTGGT	CAGGTACCTAGAAAAAGAAATAGTTTTCATAGAGC	4927
Qy	5485	TTCTCTGATG	ACTATPAGCTGCCT	CTCTGAATCAAGTCTTCTCATTTTCAGGTGCCACGGTCTG	5544
Db	4928	TTCTCTGAGG	ACTACAGCTGTCTT	CTAAATCAGGCTTCTCACTTTTAGGTGCCACGGTCTG	4987
Qy	5545	CAGATGATG	AGCGAAGACATCCT	GTCTCTGTCTTTCTGTGGGGCTATATACTATGTTCTC	5604
Db	4988	CAGATGATG	AGCGAAGACATCCT	GTCTCTGTCTTTCTGTGGGGCCATCCTCTGTGTTCTC	5047
Qy	5605	AGAACTATTT	CTCTGCCAGGAAAT	TGTGAACGGGGAAAGAGGTGTGGAGCTTGCAATTTTTCACG	5664
Db	5048	AGAACATCT	GTGTGCCAAGNA	NTAGTGAATGGGGAGAGGTTTGGAGCGTGGTTTTTCATG	5107
Qy	5665	CATTTCTCT	GTGAGCGCGAGTCT	CGAATTTTCTTAAAAATCAGAGAAATCCGAGTGTGTCTC	5724
Db	5108	CGCTTCAAT	TGTGTGTCTGAGTCT	CGAATTTTCTTAAAAATCCGAGAAATCAGGGTGTGTCTC	5167
Qy	5725	TGGTTGAAGT	TAAGCCAGAGGCTGT	GCCTATCCAGCTCCTTACTTGGATGGAATATGGAG	5784
Db	5168	TGGTGAAGAA	GAAGCCAGAGGCTGT	GCCTATCCAGCCCTTACTTGGATGGAATATGGAG	5227
Qy	5785	AAACAGACCC	TGGCTGGAAGGGG	CAACCCCTTTCATTTTATCTCGTGAGCGGTATCGGA	5844
Db	5228	AAACAGACCC	CAGGGCTAAAGAGAG	GAACCCACTTCAATTTATCTCGGGAGCGGTATCGGA	5287
Qy	5845	AGCTCCATTT	TGGTCTGGCAACA	CACTGCATTTATAGAGAGATTTGTAGAGGCCAAGAGA	5904
Db	5288	AGCTGCAATTT	TGGTCTGGCAACAG	CACACTGCAATTATAGAGAGATTTGTCTGGAGCCAGAGA	5347
Qy	5905	CTAATCAGAT	TGTTATTTGGATT	CAACTGGGCAGTTTACTGTGAGCTCCAACCTCTGCCTCAAG	5964
Db	5348	CTAATCAGAT	TGCTATTTTGGATT	TAACTGGGCAGTTTACTCTGAGCTTCAAGTTCTGCCTCAAG	5407
Qy	5965	ACAATCAAA	TGACGACAGTAGT	AAAGGCTGATTTCAAAATTTATGGAATACTTTCTGAGG	6024
Db	5408	ACAATCATG	ATGATGACATCA	TAAATAAAGACTGATCTAAATTTCTAGAGAACTTTCTGAGG	5467
Qy	6025	GCTGGGAAAG	TATTTGAGGCTTTT	TGCTCCATGTCGAGGTCACTTACATCAATAAAT	6084
Db	5468	ACGGGGGAAG	TATTTGAGGCTTTT	TGATCATGTCATGTCAGATTCACACATTAATAAAT	5527
Qy	6085	ATTTCTTAAT	TGGAGTATTTGCTTT	CAATTAGCAAAACATATGCTTTCACAGGAAAAA-AGGAC	6143
Db	5528	ATTTCTTAAT	TGGAGTATTTGCTTT	CAATTAACAACATTAGCTTTCAGGGGAAAAACAGAC	5587
Qy	6144	ATAGATCAAT	CTCTG-----	-----TTT	6160
Db	5588	ATAGATTAA	TGTTTTTATGTTCT	AGAACATGCTGTGTTTCATCCAAGTGCTAT	5647
Qy	6161	ATGTCTAGT	ATTTTTCAGGNAAT	TATTTCCCTTCATAATTTTGTCTCATTTTCAATTTATTTT	6220
Db	5648	TTCTGTCTAA	TATTTTTCAGGAAAC	TCTCTTTCCCTTCATAACTGTCTAGTTTCATTTTCATAT	5707
Qy	6221	CATCCACTCT	GGTAGATGAAGT	CACGTCMAACAGTGTAGACATTTTATGTGTTGGTTTAAC	6280
Db	5708	CACCCACTCT	GGTTAATGAGGT	CAATTAAGCATTTTGTGNCATTTTCTCCATCTGGCTAAC	5767
Qy	6281	TCTTCTGCA	ATTTTGTATTGTTGG	TTTT	6307
Db	5768	ATCTCTGC	ACCTTTGATTGTTGG	TTTT	5794

RESULT 6	
AR121463	
LOCUS	
DEFINITION	

linear PAT 16-MAY-2001



[illegible]







FEATURES	source	Location/Qualifiers
1..6395	/organism="unidentified"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32644"	
ORIGIN		
Query Match	70.3%;	Score 4435.4; DB 6; Length 6395;
Best Local Similarity	86.3%;	Pred. No. 0;
Matches 4997; Conservative	0;	Mismatches 721; Indels 69; Gaps 6;
Qy	590	TTCAGGGGCGCTCGTAAAGTGTGCTCCCTGTCTC-TCGGACCGGCCACAGGTTTCGGCT 648
Db	8	TTCAGGGGCGCTCGTAAAGTGTGCTCCCTGTGCGCTCGGCGCGGCCACAGGTTTCGGCT 67
Qy	649	TGCCTCTGCGCGGGGTCGGCAACTGCAAGCGGTCAAGTTTCCCTCAAGATGCGGACGAGG 708
Db	68	AGCTGGCGCGCGGGGTCGGGAACTGCGGGGCTTCGTTTCCCTTAAGATGGCGGACGAAG 127
Qy	709	AGGCTGGAGGTACTCAGAGGATGGAAATCAGCGGGAGTTACCCAGACCCCTCAGGGTC 768
Db	128	AGATGGACGGCGCGAGAGGATGGACGTGACGCCGGAGCTCCCTTGCCCCCGCAGCGGC 187
Qy	769	TGGCATCTTGTGTGGGATCAGCAAGTTGATTTTATATATCTGCTTCTTGATCATTTGGCAC 828
Db	188	CGGCATCGTGTGGGATCAGCAAGTTGATTTCTATCTACTGCTTCTTACATCATTTGGCAC 247
Qy	829	AATTGGTGCACAGAAATTTACTTTCTGCTGAATGGACCCAGACTTCGMAAGCAGGAGGAA 888
Db	248	ANTTAGTGCCAGAAATTTATTTTGTCTGAGATGGACCCAGATTTGGAAAAGCAAGAAGA 307
Qy	889	GTGTACAAATGTCAAATTTACTCCACTCGGTAATGGTACTTATTTGGAGAAATCCAGATA 948
Db	308	GTGTACAGATGTCAAATCTCACTCCTTTGGAGTGTACTTATTTGGAGAGGATCOGATA 367
Qy	949	TTTGCTTAGAGAAATTGAAGCACAGTGGAGCATTTTCAGCTTTGTGGAGGGTTTTCAAAA 1008
Db	368	TTTGCTTAGAGAAATTAAAAACAGTGGAGCGTTCCAGTTGTGTGGGAAGGTTTTCAAAA 427
Qy	1009	GTGCAGAGACAACTTATTTCTTGCAAGGATTTGTGCAATTTGATCCAACTGTGTACTCTGTA 1068
Db	428	GTGCAGAAACAACATATTTCTGTAGGGAATGTGCAATTTGATCCAACTGTGTCTGTGTA 487
Qy	1069	TGGACTGCTTCAGGACAGTGTTCATAAAAATCATCGTTTACAGATGCATCTTCTACTG 1128
Db	488	TGGACTGCTTCCAAAGTAGTGTTCATAAAAACCATCGTTACAGATGCATCTTCTACTG 547
Qy	1129	GAGAGGGTTCGTGACTGTGTGGACACAGAGGCATGGAAAACCTGGGCCCTTTTGTGTAA 1188
Db	548	GAGGGGGCTTCTGTGACTGTGTGGACACAGAAAGCGTGGAAAACCTGGGCCCTTTTGTGTGG 607
Qy	1189	ATCATGAACCTGGAGAGCAGTACTATTAAGAGATTCACGCTGTCCGTGTGATGATCAAG 1248
Db	608	ATCACGAGCTTGGAGAGCAGTACTACAAAGAGAGCTTACATTTGCCCAATTTGAATGAAG 667
Qy	1249	AGGTAAATTGTCCTCAAGCCAGGAAAAATATTTCTTCAGTGTATAAAATATGTCTAGAAATGA 1308
Db	668	AGGTGATTGCTCAAGCCAGGAGATATTTCCCTTCGGTGATATAAATCATTTGTAGAAATGA 727
Qy	1309	CTATATCGGAGAGGAAAAAGAACTGCTCTGAACTCCAGATTAAGGGAGAAAAATGAAA 1368
Db	728	CTATATCGGAGAGAAAAAGGAAATTGCTCTCGAACTGCAGATAAGGGAGAAAAATGAAC 787
Qy	1369	GATACTATTGTGTCCTTTTCAATGATGAACNCCATTATATATGACCACTCATATACAGCC 1428
Db	788	GATACTATTGTGTCCTTTTCAACGATGAGCAACCAATTCGTATGATCATGTGTCTACAGTC 847
Qy	1429	TACAAAGAGCTCTTGACTGTGTGAGCTTCGACAGGGCCAGTTGTGATATACCACTGCCATTGACA 1488
Db	848	TGCAGAGGCTCTAGATTGGAGGTTTGCAGAGGCACAGCTTGCAACGACTGCCATTCGACA 907
Qy	1489	AAGAGGGTCGTGGGCTGTTAAAGCGGGAGCTTATGCTGCTTGTCCAGGAAGCAAGGAAG 1548



Db 4208 TTGGAGTTGCAAGAGGGCTACTGCCCTCAGGTCTGATACAAAACATCTGGCTCGGC 4267  
Qy 4837 TTCTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACATGCTCTTCTGTCTATAG 4896  
Db 4268 TCCTGTGAGTTATTCTTCTTAACCTGCAATCAGAAAATACACAGGCGCTTCTGTCTGTGG 4327  
Qy 4897 ATCTGTTTCATGTTTTGGTGGGTGCTGTGTAGCAATCCCATCTCTGTATTTGGGATGACC 4956  
Db 4328 ATCTCTTCATGTTCTGTGTGGCGCAGTCTTAGCGTTCCCATCTCTGTATTTGGGATGACA 4387  
Qy 4957 CTGTTGATCTGAGCGCTTCTTCTCAGTTAGTTCTTCTCTATAACACCTCTTATCTCTTCCATT 5016  
Db 4388 CGTGGATCTGAGCGCTGCGCACTTAGTTCTTCAATATACCACTCTATCTCTTCCATC 4447  
Qy 5017 TGATCACCATGSCACATGCTTTCAGATACTACTTACAGTAGACA -----CAG 5064  
Db 4448 TGATCACCATGSCGACATGCTTTCAGATACTCTTACAAACAGATACAGATCTGTCTCCAG 4507  
Qy 5065 GCCTACCCCTTGTCTCAGGTTCAAGAGACAGTGAAGAGGCTCAATTCGGATCTTCTTCT 5124  
Db 4508 GCGCCGCTTGTCTGAGGCTGAAGAGGATAGTGAAGAGGCTCGCTGTGCACTCTGCTTCT 4567  
Qy 5125 TTGCAAGAAATTTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCTTGGCTGTATT 5184  
Db 4568 TTGTGGAAGTGTGCGAGCACAGACGCGCTCCTGCGGTGGGTGCTCCCGGCTGGTACC 4627  
Qy 5185 TGTGGGTCTCACTGAAGAGTGCATCACCCCTTATCTTCTGCTGCTGCTGATTTTCTTCC 5244  
Db 4628 TGTGGCTCTCCCTGAGGAAGGCACTCACCCCTTACTCTCGCTGTGCTGCACTGCTTCTCC 4687  
Qy 5245 ACTATTTACTTGGGTAACTCGCGCTGAGGAACGTGCATACCAATTTCTGCAAGAGGAGT 5304  
Db 4688 ACTATTTACTTGGAGTAGCTCGCGCTGAAGAACTGTTTGCCAAATCTGCTGAAGGAGAAT 4747  
Qy 5305 ACAGTGCACTGTGAGTATCTATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAAT 5364  
Db 4748 TCAGTGCACTGTGAGTATCTATCTTTACCCACAAATTTGTTCTGCTTCTTCCAGGAAT 4807  
Qy 5365 ATTGGGATCTGTAAAGGCCCTTGTCTCCAGAGGTGTGTGAGATCTCTGCTTACTCAAT 5424  
Db 4808 ATTGGGATACATTAAGGCCCTTACTACAGAGGTGTGTGAGATCTCTGCTTACTCAAT 4867  
Qy 5425 GTTTGAAGCAAAAACACCGTGTGTCAGGTACCTTAGAAAAGAAATAGTTGTATAGAGC 5484  
Db 4868 CTTTGAAGCAGAAAAGTGTGTGTGTCAGGTACCTTAGAAAAGAAATAGTTGTATAGAGC 4927  
Qy 5485 TTCTGTAGTACTATAGTGCCTTCTGATCAAGCTTCTCATTTCAAGTGCCTTACTCAAT 5544  
Db 4928 TTCTGTAGGACTACAGCTGTCTTCTTAATCAGGCTTCTCACTTTAGGTGTCCACGGTCTG 4987  
Qy 5545 CAGATGATGAGCAAGATCTGTCTCTGCTTCTGTTGGGCTATATCTATGTTCTC 5604  
Db 4988 CAGATGATGAGCAAGATCTGTCTCTGCTTCTGTTGGGCTATCTGTTGTTCTC 5047  
Qy 5605 AGAATTTGTGTCAGGAAATTTGAAAGGGGAGAGGTTGAGGCTTGTCAATTTTTCAG 5664  
Db 5048 AGAATCTGTTGCCAAGAAATAGTGAATGGGAGAGGTTGAGGCGTGTGCTTTCATG 5107  
Qy 5665 CACTTCACTGTGAGCGGAGTCTGCAATTTCTTAATAATCGAATGCGAGTGTGCTC 5724  
Db 5108 CGCTTCATTTGTGTGCTGAGTCTGCAATTTCTTAATAATCGAATGCGAGGTTGCTC 5167  
Qy 5725 TGGTTGAAGGTAAAGCCAGAGGCTGTGCTTACCTACAGCTCTTACTTGGATGAATATGAG 5784  
Db 5168 TGGTGAAGGAAAGCCAGAGGCTGTGCTTACCCAGGCTTACTTGGATGAATATGAG 5227  
Qy 5785 AAACAGACCTCGGCTGAAGAGGGGCAACCCCTTCTTATCTGCTGAGCGGTATCGGA 5844  
Db 5228 AAACAGACCCAGGCTTAAGAGAGGAAACCCACTTCAATTTCTCGGAGCGGTATCGGA 5287  
Qy 5845 AGCTCCATTTGTTGTGGCAACAACTGCAATATAGAGAGATTGCTAGGAGCCAGAGA 5904

Db 5288 AGCTGCATTTGGTCTGGCAACAGCACTGCATTTATAGAGAGATTGCTCGGAGCCAGGAGA 5347  
Qy 5905 CTAATCAGATGTTATTTGGATTCAATCGCGAGTTACTGTGAGCTCCAATCTGCTCTCAAG 5964  
Db 5348 CTAATCAGATGTTATTTGGATTAACTTAACTGGCAGTTACTCTGAGCTTCAGTTCTGCTCAAG 5407  
Qy 5965 ACATCACAATGAGCAGAGTAGTAGTAAGGCTGTGATTTCAAAATATGGAACCTTCTGAGG 6024  
Db 5408 ACAATCATGAGTGACATCAATAATAAGACTGATCTAAATTTCTAGAGAACTTTCTGAGG 5467  
Qy 6025 GCTGGAAAGTATTTGGAGGGTCTTTTGTCTCAATGTCAGGTTCACTTACATCAATAAAT 6084  
Db 5468 ACGGGGGAGTATTTGGAGGGTCTTTTGTATCAATGTCAGATTCACACATTAATAAAT 5527  
Qy 6085 ATTTCTTAAATGGAGTATTTGCTTCAATTAGCAACATATGCTTCAAGAAAAA-AGGAC 6143  
Db 5528 ATTCCTTAATGGAATATTTGCTTCAATTAACAATAAGCTTCAAGGGAAAAACAAGAC 5587  
Qy 6144 ATAGATCAATCTGT-----TTT 6160  
Db 5588 ATAGATTAATGTTTATGTTCTAGAACACTAAAGAAATGCTTGTTCATCCAGTGTCTAT 5647  
Qy 6161 ATGTGCTAGTATTTCCAGGAATTTATCCCTTTCATATTTGTCTCAATTTCTCAATTTATTT 6220  
Db 5648 TTCTGCTAATATTTCCAGAAACTCTTCCCTTCATTAACCTGCTAGTTCATTTCATAT 5707  
Qy 6221 CATCCACTGTTAGTAGTAAGTCAAGTCAACAGTTGTAGACATTTATATGTTGTGTTAAC 6280  
Db 5708 CACCACCTGTTAATGAGGTCAATTAAGCATTTGTGACATTTCTCCATCTGCTAAC 5767  
Qy 6281 TCCTTCTGCAATTTGTTATTTGTTGTTT 6307  
Db 5768 ATCTCTGCACCTTGTATTTGTTGTTT 5794

RESULT 8  
AF061555 6395 bp mRNA linear ROD 10-JUL-1998  
LOCUS Mus musculus ubiquitin-protein ligase E3-alpha (Ubr1) mRNA,  
DEFINITION complete cds.  
ACCESSION AF061555  
VERSION AF061555.1 GI:3170886  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 6395)  
AUTHORS Kwon, Y.T. and Varshavsky, A.  
TITLE Direct Submission  
JOURNAL Submitted (24-APR-1998) Division of Biology, 147-75, California  
Institute of Technology, 1200 E. California Bl., Pasadena, CA  
91125, USA  
FEATURES  
source  
1. .6395  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/chromosome="2"  
/cell\_line="MEL-C19"  
1. .6395  
/gene="Ubr1"  
1. .114  
/gene="Ubr1"  
115. .5388  
gene  
5'UTR  
CDS







Qy	4240	GCATTCAGTGTGACCTTTTGTGACTTGGAAAGTGGAGAAATATCTTTGGCCCTCTTTGGCAAT	4299
Db	3668	GCATTCACGTAGACCTGTTTGACCTGGAGAGCGCGAGTAGTACCTATGATGCCGCTCTGCAAGT	3727
Qy	4300	CTCTGTGCAATACGTGTGATCCCCATTTATCTTTTGCNAACCTCAAAAGATATAACAGTGCAGA	4359
Db	3728	CTCTCTGCNAACACTGTTCATCCCCATCATCCCTTTGACGCGCGAGAGATCAACAGATGAGA	3787
Qy	4360	ATGCAGATGCTCTTGCTCAACTTTTGACCTTGGCAGCGGTGGATACAGACTGTCTTGCCCA	4419
Db	3788	ATGCGGAGGCTCTTGCTCAACTTTTGACCTTTGGCCCGGTGGATACAGACTGTCTTGCCCA	3847
Qy	4420	GAATATCAGGTTATATATATATAGACATGCTATAAAGSAGAAACC---CAATTCCTATTTTCT	4476
Db	3848	GAATATCGGGTTATAATATAAGCATGCTATAAGSAGAAAGCCCAAGCAGTTCCTGTCTTGT	3907
Qy	4477	TTAATCAAGGAATGGGAGATTTCTACTTTTGGAGTTCATTTCCATCTCTGAGTTTTTGGCGTTG	4536
Db	3908	TTAATCAAGGAATGGGCGGTTTCAACTTTTCAAGTTTCAATCCATCTCTGAGTTTTTGGAGTTC	3967
Qy	4537	AGTCTTCGATTTAAATATTCAAAATAGCATCAAGGAATGGTTATTCTCTTTGCCACAAACA	4596
Db	3968	AGTCTTCGGTGAATATTCAAAATAGTATCAAGGAATGGTCATTCTCTTCGCCACAAACA	4027
Qy	4597	TTTATAGAAATTGGATTGAAGTGCCACCTGATGAAGGAGATCTTCGAGTCCCATGCTGA	4656
Db	4028	TTTTACAGAAATTGGCTTGAAAGTGCTCTCTGATGAATCTAGACCCACGAGTGCCCATGATGA	4087
Qy	4657	CTTGGAGACCTTGGCGCTTTCATCTATCAGGCAATTTGAAATCTATTGGGAGATGAGGAA	4716
Db	4088	CCTGGAGCACGTGTGCGTTCCACCATCAGGCAATTCGAAAAACCTGTTGGGAGATGGAAGAA	4147
Qy	4717	AACCTCTGTTTGGAGCACTTCAAAATAGGCAGCATATGTCTCTGAAAGCATTAATGCAGT	4776
Db	4148	AACCTCTATTTGGAGCACTTCAAAATAGACAGCATAGCGTCTGAGGCCCTTAATGCAGT	4207
Qy	4777	TTGCAGTTTGCAACAGAGGATTAACCTGTCTCTCAGGTCTTGATACAGAAACATCTGTTGCTC	4836
Db	4208	TTGCAGTTTGCAACAGAGGCTACCTGCGCTCAGGTCTTGATACAGAAACATCTGCTCGGC	4267
Qy	4837	TTCTATCAGTTGTTCTTCTAAATATAAAATCAGAGATACACCATGCTCTGTCTATATAG	4896
Db	4268	TCCTGTCAAGTTATTCTTCTTAACCTGCAATCAGAAAAATACACCAAGGCTCTCTGTCTGTG	4327
Qy	4897	ATCTGTTTCAATGTTTGTGGTGGTCTGTGTTAGCATTCCTGATCTTGTATTTGGATGACC	4956
Db	4328	ATCTCTTCCATGTTCTGGTGGCGCAGTCTTAGGGTTCCCATCTTGATTTGGATGACA	4387
Qy	4957	CTGTTGATCTGCAGCCCTTCTTCAGTTAGTTCTCTCTATAACCACTTTATCTCTTCCATT	5016
Db	4388	CCGTGGATCTGCAGCCGCTCGCCACTTAGTTCTTTCATATAACCACTCTATCTCTTCCATC	4447
Qy	5017	TGATCAACATGGCAACAATGCTTCAGATATCTTAATCAGTAGACA-----CAG	5064
Db	4448	TGATCAACATGGCGCACATGCTTCAGATATCTCTCTTAACAGAGATACAGATCTGTCTCCAG	4507
Qy	5065	GCCTTACCCCTTGCTCAGGTTCAAGAACACAGTGAAGAGGCTCATTCGCCATCTCTTTCT	5124
Db	4508	GGCCGCGCTTGCTCGAGGGTGAAGAGATAGTGAAGGAGGCTCGCGTGTGCATCTGCTTTCT	4567
Qy	5125	TTGCAGAAATTTCTCAATATACAAAGTGGCTCCATTTGGGTGTGATTTCTCTGGCTGGTATT	5184
Db	4568	TTGTGGAGTGTGCGAGCAACACAGCCGCTCACTGGGTGGGTGCTCCCGGCTGGTACC	4627
Qy	5185	TGTGGGTCTCATGAAAGATGGCAATCACCCCTTATCTTGTGTGTGTGCAATTTGTTTTTCC	5244
Db	4628	TGTGGCTCTCCCTGAGGAAAGGCAATCACCCCTTACCTCCGCTGTGTCTGCACCTGTCTTTCC	4687
Qy	5245	ACTATTTACTTGGGTTAACTCCGCTCAGGAACTGCATACCAATTCCTGCAAGAGGAGAT	5304
Db	4688	ACTATTTACTTGGAGTAGTCTCCGCTGAAAGAACTGTTTGGCAATTCCTGCTGAAGAGGAT	4747
Qy	5305	ACAGTGCACTCTGTAGCTATCTATCTTTTACCTAACAAATTTGTTTGGTCTCTCTCCAGGAA	5364

CQ728565 LOCUS CQ728565 4529 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 14499 from Patent WO02068579.  
ACCESSION CQ728565  
VERSION CQ728565.1 GI:42297720  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 14499 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
Source 1..4529  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 55.1%; Score 3473.6; DB 6; Length 4529;  
Best Local Similarity 82.5%; Pred. No. 0;  
Matches 4475; Conservative 0; Mismatches 24; Indels 927; Gaps 6;  
QY 550 CGGGCCGGAAGCACTCTCGAGTCGCGTCAAAACCGACCTTCAGGGCCGTCGTAAG 609  
DB 1 CGGGCCGGAAGCACTCTCGAGTCGCGTCAAAACCGACCTTCAGGGCCGTCGTAAG 60  
QY 610 TGTGCTCCCTGTCTCTCCGACCGGCCACAGGTTTCGCTTCGCTTCGCGCCGGGGTCGGC 669  
DB 61 TGTGCTCCCTGTCTCTCCGACCGGCCACAGGTTTCGCTTCGCTTCGCGCCGGGGTCGGC 120  
QY 670 AACTGCAGCGCTCAGTTTCCCTCAAGATGGCGGACGAGGCTCGAGGTACTGAGAGGA 729  
DB 121 AACTGCAGCGCTCAGTTTCCCTCAAGATGGCGGACGAGGCTCGAGGTACTGAGAGGA 180  
QY 730 TGGAAATCAGCGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGTGGGATCAGC 789  
DB 181 TGGAAATCAGCGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGTGGGATCAGC 240  
QY 790 AAGTTGATTTTATATCTGCTTTCTTGATCATTTTGGCACAATTTGGTCCAGAAATTTACT 849  
DB 241 AAGTTGATTTTATATCTGCTTTCTTGATCATTTTGGCACAATTTGGTCCAGAAATTTACT 300  
QY 850 TTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTAACAATGTCAATATTCA 909  
DB 301 TTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTAACAATGTCAATATTCA 360  
QY 910 CTCCACTGGATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAGC 969  
DB 361 CTCCACTGGATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAGC 420  
QY 970 ACAGTGGAGCATTTGAGTGGGGGGTTTTCAAAAGTGGAGAGCAACCTATTCTT 1029  
DB 421 ACAGTGGAGCATTTGAGTGGGGGGTTTTCAAAAGTGGAGAGCAACCTATTCTT 480  
QY 1030 GCAGGATTTGCAATTTGATCCAAATGCTGCTCTGATGAGCTGCTTCAGGACAGTG 1089  
DB 481 GCAGGATTTGCAATTTGATCCAAATGCTGCTCTGATGAGCTGCTTCAGGACAGTG 540  
QY 1090 TTCAATAAATCATGTTTACAGATGCATCTTCTACTGGAGGAGGGTTCTGTGACTGTG 1149  
DB 541 TTCAATAAATCATGTTTACAGATGCATCTTCTACTGGAGGAGGGTTCTGTGACTGTG 600  
QY 1150 GAGACACAGGCGCATGGAAATCTGCCCTTTTGTGTAATCATGAACTCGAAGAGCAG 1209  
DB 601 GAGACACAGGCGCATGGAAATCTGCCCTTTTGTGTAATCATGAACTCGAAGAGCAG 660  
QY 1210 GTACTATAAAGAGAAATTCACGCTGTCCGTTTGAATGAAGAGGTAAATTTGTCAGCCAGGA 1269

DB 661 GTACTATAAAGAGAAATTCACGCTGTCCGTTTGAATGAAGAGGTAAATTTCTCAAGCCAGA 720  
QY 1270 AAATAATTTCTTCCAGTGAATAAATATGTCGTAGAAATGACTATATGCGAAGAGGAAAAAG 1329  
DB 721 AAATAATTTCTTCCAGTGAATAAATATGTCGTAGAAATGACTATATGCGAAGAGGAAAAAG 780  
QY 1330 AACTGCTCTCTGAACTCCAGATAAGGGAGAAAAATAGAAAGATATATTTGTGTCTTTTCA 1389  
DB 781 AACTGCTCTCTGAACTCCAGATAAGGGAGAAAAATAGAAAGATATATTTGTGTCTTTTCA 840  
QY 1390 ATGATGAACACCATTTATATGACCAACGTCATATACAGCTTACAAAGAGCTCTTTGACTGTG 1449  
DB 841 ATGATGAACACCATTTATATGACCAACGTCATATACAGCTTACAAAGAGCTCTTTGACTGTG 900  
QY 1450 AGCTCGCAGAGCGCCAGTTGCATACCACTGCTCCATTTGACAAAGAGGGTCTCGGGCTGTTA 1509  
DB 901 AGCTCGCAGAGCGCCAGTTGCATACCACTGCTCCATTTGACAAAGAGGGTCTCGGGCTGTTA 960  
QY 1510 AAGCGGAGCTTATGCTGCTTGCAGGAAGCAAAAGGAAGATATAAAGAGTCAATTCAGAAA 1569  
DB 961 AAGCGGAGCTTATGCTGCTTGCAGGAAGCAAAAGGAAGATATAAAGAGTCAATTCAGAAA 1020  
QY 1570 ATGTCTCTCAACATCCACTTCATGTAAGATTTACACTCAGAGATTTATGCTCATCAGA 1629  
DB 1021 ATGTCTCTCAACATCCACTTCATGTAAGATTTATACACTCAGAGATTTATGCTCATCAGA 1080  
QY 1630 AATTGCTTTGCGTCTTGGTTCTTGGATGAACAAAATTTATGAGCTATTTCAAGTGAATTTA 1689  
DB 1081 AATTGCTTTGCGTCTTGGTTCTTGGATGAACAAAATTTATGAGCTATTTCAAGTGAATTTA 1140  
QY 1690 GGCAAGATCTTTTTCAGACATGCTTTAGAGAAAGAACTGACTCGGAGAAATCCCTGTCTCA 1749  
DB 1141 GGCAAGATCTTTTTCAGACATGCTTTAGAGAAAGAACTGACTCGGAGAAATCCCTGTCTCA 1200  
QY 1750 TAACAGGTTAATGCTTTGGATGCAAGCTTTTATAAAGTGGCGGTAAAGATCCTTCAATG 1809  
DB 1201 TAACAGGTTAATGCTTTGGATGCAAGCTTTTATAAAGTGGCGGTAAAGATCCTTCAATG 1260  
QY 1810 AATTGATCTTCAGCAGATTTTATGAGATGGAATACAAAATACTCTTCTGCTATGGAAT 1869  
DB 1261 AATTGATCTTCAGCAGATTTTATGAGATGGAATACAAAATACTCTTCTGCTATGGAAT 1320  
QY 1870 TTGTGAAGTATTAATAAACAATCGCAGAAAGAAATATACAGTGAATCATGACAGAAATA 1929  
DB 1321 TTGTGAAGTATTAATAAACAATCGCAGAAAGAAATATACAGTGAATCATGACAGAAATA 1380  
QY 1930 TCTCTATAAATGCACTTTTCACTGATGTTTACCTGCTTCTACTCTGCTGCTGACATCTTA 1989  
DB 1381 TCTCTATAAATGCACTTTTCACTGATGTTTACCTGCTTCTACTCTGCTGCTGACATCTTA 1440  
QY 1990 TTGAAGAGCAGAAATGTTTATCTGCTCAATCTGAAACTCTGCTAGAAAGTTTCTGCTGAGT 2049  
DB 1441 TTGAAGAGCAGAAATGTTTATCTGCTCAATCTGAAACTCTGCTAGAAAGTTTCTGCTGAGT 1500  
QY 2050 ACTTGGCAGGAAACAATAAATTTCAACTTCCAGGGTTATAGCCAGGACAAAATTTGGGAAG 2109  
DB 1501 ACTTGGCAGGAAACAATAAATTTCAACTTCCAGGGTTATAGCCAGGACAAAATTTGGGAAG 1560  
QY 2110 TATATGAGTAAATATGTAACCTAAAGTATATCTCTGATCAGCAAAACCCCAATATGGAACAG 2169  
DB 1561 TATATGAGTAAATATGTAACCTAAAGTATATCTCTGATCAGCAAAACCCCAATATGGAACAG 1620  
QY 2170 AAGATTAAGAAATGCAAGTTCTTCAAGGTTTTCAGATCTTTTTCAGAAATTTCTTACCTGTA 2229  
DB 1621 AAGATTAAGAAATGCAAGTTCTTCAAGGTTTTCAGATCTTTTTCAGAAATTTCTTACCTGTA 1680  
QY 2230 TGCAAGGAAATGGAAGAAATCCGAAAGACAGGTTGGGCAACCAATTTGAAGTGGATCCTGAT 2289  
DB 1681 TGCAAGGAAATGGAAGAAATCCGAAAGACAGGTTGGGCAACCAATTTGAAGTGGATCCTGAT 1740  
QY 2290 GGGAGGCTGCAATTTGCTATACAGATGCAATTTGAAGAAATTTTATCTCATGTTCCAAAGAT 2349

Db 1741 GGGAGGCTGCCATTGCTATACAGATGCAATTTGAAGAAATATTTTACTCATGTTCGAAGAGT 1800  
Qy 2250 GGTGTGCTTGATGAAGAACTCTTACTTGTGGCTTATAAAGAAATGTCACAAAGCTGTGA 2409  
Db 1801 GGTGTGCTTGATGAAGAACTCTTACTTGTGGCTTATAAAGAAATGTCACAAAGCTGTGA 1860  
Qy 2410 TGAGGTGAGTACCAAGTTTTCATATCTAGTAGCAAGACAGTAGTACAATCGTGTGGACATA 2469  
Db 1861 TGAGGTGAGTACCAAGTTTTCATATCTAGTAGCAAGACAGTAGTACAATCGTGTGGACATA 1920  
Qy 2470 GTTTGGAAAACAAAGCTTCAAGATCTGAGGATCTTTGTAAGCATACATCTGCCACTCT 2529  
Db 1921 GTTTGGAAAACAAAGCTTCAAGATCTGAGGATCTTTGTAAGCATACATCTGCCACTCT 1980  
Qy 2530 CTAGGACCCCTTGCTGCTTTCATGTACATGCTTTAAGCAGGCTGGGTCTGTTTCAAGACTGC 2589  
Db 1981 CTAGGACCCCTTGCTGCTTTCATGTACATGCTTTAAGCAGGCTGGGTCTGTTTCAAGACTGC 2040  
Qy 2590 ATGAATTTGTCTTTTTCAGGACTTTCAAGTAGAGGTACTAGTGGAAATATCTTTACGTT 2649  
Db 2041 ATGAATTTGTCTTTTTCAGGACTTTCAAGTAGAGGTACTAGTGGAAATATCTTTACGTT 2100  
Qy 2650 GTCTGGTGTGGTTCGCCAGGTGCTGCTGAGATGTGGCGAAGAAATGGAAGTCTCTTA 2709  
Db 2101 GTCTGGTGTGGTTCGCCAGGTGCTGCTGAGATGTGGCGAAGAAATGGAAGTCTCTTA 2160  
Qy 2710 TTAGCCAGGTGTTTATTAACAGATGTTAAGTGCAGAGAAATGATGATAAAGATA 2769  
Db 2161 TTAGCCAGGTGTTTATTAACAGATGTTAAGTGCAGAGAAATGATGATAAAGATA 2220  
Qy 2770 TCATCATGCTT-----CAGATTGGTGCATCTTTAA 2799  
Db 2221 TCATCATGCTTCAGAAAGAAATATGATGCTGGATTTCTTTCAAGATTGGTGCATCTTTAA 2280  
Qy 2800 TGGATCCCAATAAGTTCTTTGTTACTGTGTACTTCAGAGGTATGAATTCGCCAGGCTTTTA 2859  
Db 2281 TGGATCCCAATAAGTTCTTTGTTACTGTGTACTTCAGAGGTATGAATTCGCCAGGCTTTTA 2340  
Qy 2860 ACAAGACATATCTACAAAGACAGGATTTGATTAACAATATTAATACATTAAGAG 2919  
Db 2341 ACAAGACATATCTACAAAGACAGGATTTGATTAACAATATTAATACATTAAGAG 2400  
Qy 2920 AAATGCTTTCAGTCTCATCTATATTTGTTGGGTGAGCGTTATGTACTCGAGTGGGAATG 2979  
Db 2401 AAATGCTTTCAGTCTCATCTATATTTG----- 2429  
Qy 2980 TGACCAAGAGAGGTGCAAAATGAGAGAAATCAITTCATCTGCTTTGCAATGAACCCATGC 3039  
Db 2430 ----- 2429  
Qy 3040 CACACAGTGCCATTGCCAAAATTTTACCTGAGAAATGAATAATGAATGCTGTAGAGA 3099  
Db 2430 ----- 2429  
Qy 3100 ATGTCAATAACAAAGTGGCCACATTTAAGAAACAGGTGTATCAGGCCATGAGTTTATG 3159  
Db 2430 ----- 2429  
Qy 3160 AACTAAAGATGAATCACTGAAGAGCTTCAATATGTACTTTTATCAITTAATCTCCAAACCC 3219  
Db 2430 ----- 2429  
Qy 3220 AGCATAGCAAGGCTGAACATATGCAAGAAAGAGAGAAACAAAGAAATGAAG 3279  
Db 2430 -----G 2430  
Qy 3280 CATTCGCCCCACCACTCCTGAAATTCCTGCTGCTTTTCAGCAAGAGTATTAACCTTC 3339  
Db 2431 CATTCGCCCCACCACTCCTGAAATTCCTGCTGCTTTTCAGCAAGAGTATTAACCTTC 2490  
Qy 3340 TCAACTGTGATCATGATGTACATTTCTCAGGACCGTATTTTTCAGGCGGCAATAGACAG 3399  
Db 2491 TCAACTGTGATCATGATGTACATTTCTCAGGACCGTATTTTTCAGGCGGCAATAGACAG 2550

Qy 3400 ATTCTAACTTGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGSCATTTGGTT 3459  
Db 2551 ATTCTAACTTGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGSCATTTGGTT 2610  
Qy 3460 TACTAGAAGAGAGCAACAGCTTCAAAAGCTCCTGAAGAGAAAGTAACTTTGACATTTT 3519  
Db 2611 TACTAGAAGAGAGCAACAGCTTCAAAAGCTCCTGAAGAGAAAGTAACTTTGACATTTT 2670  
Qy 3520 ATCATAGGCTTCAAGATTTGGGAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAGAC 3579  
Db 2671 ATCATAGGCTTCAAGATTTGGGAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAGAC 2730  
Qy 3580 TCAAGGAATTTCCCGAGTTAGNAGGCCAGAGGACATGATAAAGCTGATATCTTCAGATGT 3639  
Db 2731 TCAAGGAATTTCCCGAGTTAGNAGGCCAGAGGACATGATAAAGCTGATATCTTCAGATGT 2790  
Qy 3640 TTGACACAGTGAAGCGAATTAAGAGAAAAATCTTGTGTTAAATTTGTAGCAACCAATCAGAT 3699  
Db 2791 TTGACACAGTGAAGCGAATTAAGAGAAAAATCTTGTGTTAAATTTGTAGCAACCAATCAGAT 2850  
Qy 3700 CGGAATCTTAAAGATGATGAGATTACTCATGATAAAGAAAAAGCAGAAAGAA 3759  
Db 2851 CGGAATCTTAAAGATGATGAGATTACTCATGATAAAGAAAAAGCAGAAAGAA 2910  
Qy 3760 AAGCTGAAGCTGTAGGCTTACATCGCAGAGATCATGGCTCAGATGCTTCGCCTTACAGA 3819  
Db 2911 AAGCTGAAGCTGTAGGCTTACATCGCAGAGATCATGGCTCAGATGCTTCGCCTTACAGA 2970  
Qy 3820 AAAATCTTCAITGAACTCAATAACTCATGTATGACAAATCATACATCAAGAAATGCTTCGGAAAG 3879  
Db 2971 AAAATCTTCAITGAACTCAATAACTCATGTATGACAAATCATACATCAAGAAATGCTTCGGAAAG 3030  
Qy 3880 AAGATTCATTTAGAGAGAGAGAGACCCAGCAGCTCAGTACTCTTGAATTTGCTT 3939  
Db 3031 AAGATTCATTTAGAGAGAGAGAGACCCAGCAGCTCAGTACTCTTGAATTTGCTT 3090  
Qy 3940 TGGTCTCTAAACGGGTCCATCTCTTACTGAAAAGAGGTGCTGACGTGCAATCTTTGCC 3999  
Db 3091 TGGTCTCTAAACGGGTCCATCTCTTACTGAAAAGAGGTGCTGACGTGCAATCTTTGCC 3150  
Qy 4000 AAGAAAGACAGGAGGTGAAAATAGAAAATTAATCCATGTTATATTCGGCTGTGTCCAGA 4059  
Db 3151 AAGAAAGACAGGAGGTGAAAATAGAAAATTAATTCATGCTGTTATTCGGCTGTGTCCAGA 3210  
Qy 4060 AATCTACTGCTTAACTCCAGCAGAGGGGAAACCCATAGACTCTTCAGGAGAGCCCTAG 4119  
Db 3211 AATCTACTGCTTAACTCCAGCAGAGGGGAAACCCATAGAACTCTTCAGGAGAGCCCTAG 3270  
Qy 4120 ACCCACTTTTTCATGATTCAGACTTTGGCATATGGAACTTATACAGAGAGCTGTGTCTATG 4179  
Db 3271 ACCCACTTTTTCATGATTCAGACTTTGGCATATGGAACTTATACAGAGAGCTGTGTCTATG 3330  
Qy 4180 TAATGACGCAAGTGTGTCGAGAGAAATTTTGAAGCTGTGACAGTCAAGCTCTCAGCAGC 4239  
Db 3331 TAATGACGCAAGTGTGTCGAGAGAAATTTTGAAGCTGTGACAGTCAAGCTCTCAGCAGC 3390  
Qy 4240 GCATTCATGTTGACTTTTTCAGCTTGGAAAGTGAAGATATCTTTGCTCTTTTGCATTAAT 4299  
Db 3391 GCATTCATGTTGACTTTTTCAGCTTGGAAAGTGAAGATATCTTTGCTCTTTTGCATTAAT 3450  
Qy 4300 CTCTGTGCAATPACTGATCCCAATTTCTTTGCAACCTCAAAAGATAAAGTGAAG 4359  
Db 3451 CTCTGTGCAATPACTGATCCCAATTTCTTTGCAACCTCAAAAGATAAAGTGAAG----- 3505  
Qy 4360 ATGAGATGCTCTTTGCTCAACTTTTGAACCTTCGGCAAGTGGATACAGACTGTTCTGGCCA 4419  
Db 3506 ----- 3505  
Qy 4420 GAATATCAGGTTTATATATAGACATGCTAAAGAGAGAAACCCAAATTCCTATTTCTTTA 4479  
Db 3506 ----- 3505



```
QY 4480 ATCAAGGAATGGAGATTCTACTTTGGAGTTCCATTCCATCTGAGTTTGGCGTTGAGT 4539
Db 3506 ----- 3505
QY 4540 CTTGATTAATATTAATCAATAGCATCAAGGAATGTTTATTTCTCTTTGCCCAACAATTT 4599
Db 3506 ---GATTAAATATTAATAGCATCAAGGAATGTTTATTTCTCTTTGCCCAACAATTT 3561
QY 4600 ATAGAAATTTGAATGGAATGCGACCTGATGAAAGGGATCCTCGAGTCCCAATCTGAGCT 4659
Db 3562 ATAGAAATTTGAATGGAATGCGACCTGATGAAAGGGATCCTCGAGTCCCAATCTGAGCT 3621
QY 4660 GGAGCAGCTGGCTTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAAGGAAC 4719
Db 3622 GGAGCAGCTGGCTTTTCACTATCCAGGCAATTT-----GTAAG 3660
QY 4720 CTCTGTTTGGAGCACTTCAAAATAGGAGCAATAATGGTCTGAAAGCAATAATGCAAGTTG 4779
Db 3661 CTCTAAGAAACTGGTTTGGAGGTTTCAATAATGGTCTGAAAGCAATAATGCAAGTTG 3720
QY 4780 CAGTTGCAAGAGGATTAACCTGCTCAGGTCCTGATACAGAAACATCTGGTTGCTCTTC 4839
Db 3721 CAGTTGCAAGAGGATTAACCTGCTCAGGTCCTGATACAGAAACATCTGGTTGCTCTTC 3780
QY 4840 TATCAGTTGTTCTTCCCTAACATTAATCAGAAAGATACACCATGCTTCTGTCTATAGATC 4899
Db 3781 TATCAGTTGTTCTTCCCTAACATTAATCAGAAAGATACACCATGCTTCTGTCTATAGATC 3840
QY 4900 TGTTTCATGTTTGGTGGGCTGTGTTAGCAATCCCATCTCTGTTATTTGGGATGACCCCTG 4959
Db 3841 TGTTTCATGTTTGGTGGGCTGTGTTAGCAATCCCATCTCTGTTATTTGGGATGACCCCTG 3900
QY 4960 TTGATCTGCAGCCTTCTTCAGTTAGTTCTTCTATAACCACTTTATCTCTTCCATTTGA 5019
Db 3901 TTGATCTGCAGCCTTCTTCAGTTAGTTCTTCTATAACCACTTTATCTCTTCCATTTGA 3960
QY 5020 TCACATGGCACACATGCTTACAGTACTACTTACAGTACAGACAGGCTACCCCTTGCTC 5079
Db 3961 TCACATGGCACACATGCTTACAGTACTACTTACAGTACAGACAG----- 4005
QY 5080 AGGTTCAAGAGACAGTGAAGAGGCTCATTCGGCATCTTCTTTTTCGAGAAATTTCTC 5139
Db 4006 ----- 4005
QY 5140 AATATACAAGTGGCTCCATTGGGTGTGATATTCCTGGCTGTGATTTTGGGCTCTCACTGA 5199
Db 4006 ----- 4005
QY 5200 AGAATGGCATCACCCCTTATCTTGCTGTGCTGCATTTGTTTTCATTTACTTTGGGG 5259
Db 4006 ----- 4005
QY 5260 TAACCTCGCCTGAGGAATCGCATACCAATTCGAGAGGAGGATPACAGTGCACCTCTGTA 5319
Db 4006 -----ATTCTGCAAGAGGAGGATPACAGTGCACCTCTGTA 4038
QY 5320 GCTATCTATCTTTACCAAAATTTGTTCTGCTCTCTCCAGGAATATTTGGGATCTGTAA 5379
Db 4039 GCTATCTATCTTTACCAAAATTTGTTCTGCTCTCTCCAGGAATATTTGGGATCTGTAA 4098
QY 5380 GGCCCTTGCTCCAGAGGTTGTGAGATCTGCTGCTTACTAACTGTTTGAAGCAAAAAA 5439
Db 4099 GGCCCTTGCTCCAG----- 4112
QY 5440 ACACCGTGGTCAGGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTGTACTATA 5499
Db 4113 ----- 4112
QY 5500 GCTGCTCTCTGAATCAAGCTTCTCATTTACGTTGCCACGGTCTGAGATGATGAGGAA 5559
Db 4113 -----AGTGGCCACGGTCTGAGATGATGAGGAA 4143
QY 5560 AGCATCTGCTCTGCTCTTTTCTGTGGGGCTATATCTATGTTCTCAGAAACATTTGCTGCC 5619
```

```
Db 4144 AGCATCTGCTCTCTGCTCTTTTCTGTGGGGCTATATGTTCTCAGAACATTTGCTGCC 4203
QY 5620 AGGAAATTTGTAACGGGAAGAGGTTGGAGCTTGCATTTTTCACGCACCTTCACTGTGGAG 5679
Db 4204 AGGAAATTTGTAACGGGAAGAGGTTGGAGCTTGCATTTTTCACGCACCTTCACTGTGGAG 4263
QY 5680 CCGAGTCTGCATTTTCTTAAATTCAGAGAAATGCCGAGTGGTCTGTTGAGGTAAAG 5739
Db 4264 CCGAGTCTGCATTTTCTTAAATTCAGAGAAATGCCGAGTGGTCTGTTGAGGTAAAG 4323
QY 5740 CCAGAGGCTGTGCTTATCCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTTGCC 5799
Db 4324 CCAGAGGCTGTGCTTATCCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTTGCC 4383
QY 5800 TGAAGAGGGGCAACCCCTTCAITTTATCTCTGAGCGGTATCGGAAGCTCCATTTGGTCT 5859
Db 4384 TGAAGAGGGGCAACCCCTTCAITTTATCTCTGAGCGGTATCGGAAGCTCCATTTGGTCT 4443
QY 5860 GGCAACAACACTGATATAGAGAGATTCCTAGAGGCAAGAGACTAATCAGATGTTAT 5919
Db 4444 GGCAACAACACTGATATAGAGAGATTCCTAGAGGCAAGAGACTAATCAGATGTTAT 4503
QY 5920 TTGGATTCAACTGGCAGTTACTGTGA 5945
Db 4504 TTGGATTCAACTGGCAGTTACTGTGA 4529

RESULT 10
AX714232
LOCUS AX714232 3059 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 916 from Patent BP1293569.
ACCESSION AX714232
VERSION AX714232.1 GI:29889184
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J.I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
Masuho, Y.
Masuho, Y.
Full-length cDNAs
Patent: BP 1293569-A 916 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)
FEATURES
source
location/Qualifiers
1. 3059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 48.4%; Score 3055.8; DB 6; Length 3059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 680 GTGATTTTCCCTCAAGATGGCGGACGAGGCTGGAGGTACTGAGAGGATGGAATCAG 739
Db 1 GTGATTTTCCCTCAAGATGGCGGACGAGGCTGGAGGTACTGAGAGGATGGAATCAG 60
QY 740 CGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGGATCAGCAAGTTGATTT 799
Db 61 CGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGGATCAGCAAGTTGATTT 120
QY 800 TTATACGCTTTCTTGCATCATTTGGCACAATTTGGTGCAGAAATTTACTTTGCTGAAT 859
Db 121 TTATACGCTTTCTTGCATCATTTGGCACAATTTGGTGCAGAAATTTACTTTGCTGAAT 180
QY 860 GGACCCAGACTTGAAAGACGAGGAGGAAGTGTACAATGTCAATATTTCACTCCACTGGA 919
```

Db 181 GGACCAGACTTGAAAGACGAGGAAAGTGTACAAATGTCAAAATTTCACTCCACTGGA 240  
Qy 920 ATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAAGCACAGTGGAGC 979  
Db 241 ATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAAGCACAGTGGAGC 300  
Qy 980 ATTTTCAGCTTTGCGGAGGGTTTTCAAAAGTGGAGAGACCACTTATTTCTTCGACGGATTG 1039  
Db 301 ATTTTCAGCTTTGCGGAGGGTTTTCAAAAGTGGAGAGACCACTTATTTCTTCGACGGATTG 360  
Qy 1040 TGCATTTGATCCAACTGATGTACTCTGTATGGAATGCTTCCAGGACAGTGTTCATAAAA 1099  
Db 361 TGCATTTGATCCAACTGATGTACTCTGTATGGAATGCTTCCAGGACAGTGTTCATAAAA 420  
Qy 1100 TCATCGTTACAGATGCATCTTCTACTGAGGAGGGTTCTGTGACTGTGAGACACAGA 1159  
Db 421 TCATCGTTACAGATGCATCTTCTACTGAGGAGGGTTCTGTGACTGTGAGACACAGA 480  
Qy 1160 GGCAATGGAAGAACTGGCCCTTTTGTGTAAATCATGAACTTGGAAAGACAGGTACTATAAA 1219  
Db 481 GGCAATGGAAGAACTGGCCCTTTTGTGTAAATCATGAACTTGGAAAGACAGGTACTATAA 540  
Qy 1220 AGAGAAATTCACGCTGTCCGTTGAAATGAAGAGGTAAATGTCCAGCCAGGAAAAATTTTCC 1279  
Db 541 AGAGAAATTCACGCTGTCCGTTGAAATGAAGAGGTAAATGTCCAGCCAGGAAAAATTTTCC 600  
Qy 1280 TTCAAGTGAATAATATGTCGTAGAAATGCACTATATGGGAAGAGGAAAAAGAACTGCCCTCC 1339  
Db 601 TTCAAGTGAATAATATGTCGTAGAAATGCACTATATGGGAAGAGGAAAAAGAACTGCCCTCC 660  
Qy 1340 TGAATCCAGATAGAGGAGAAAAATGAAAGATACCTATTTGTGCTCTTTTCAATGATGAACA 1399  
Db 661 TGAATCCAGATAGAGGAGAAAAATGAAAGATACCTATTTGTGCTCTTTTCAATGATGAACA 720  
Qy 1400 CCATTTATATGACCACTCATATACAGCTTCAAGAGCTCTTGAAGTGTGAGCTCGCAGA 1459  
Db 721 CCATTTATATGACCACTCATATACAGCTTCAAGAGCTCTTGAAGTGTGAGCTCGCAGA 780  
Qy 1460 GGCCAGTGTGATACCACTGCATTTGACAAAGAGGGTGTGGGCTGTAAAGCGGGAGC 1519  
Db 781 GGCCAGTGTGATACCACTGCATTTGACAAAGAGGGTGTGGGCTGTAAAGCGGGAGC 840  
Qy 1520 TTATGCTCTTTGCCAGGAGCAAGAGGAGATATAAGAGTCAATTCAGAAAAATGCTCTCA 1579  
Db 841 TTATGCTCTTTGCCAGGAGCAAGAGGAGATATAAGAGTCAATTCAGAAAAATGCTCTCA 900  
Qy 1580 ACATCCACTTCAATGTAGAAGTATTAACCTCAGAGATTAAGCTCATCAGAAAAATTTGCTTT 1639  
Db 901 ACATCCACTTCAATGTAGAAGTATTAACCTCAGAGATTAAGCTCATCAGAAAAATTTGCTTT 960  
Qy 1640 GCGTCTTGGTTCTGGATGAACAAAAATTAAGCTATTTCAAGTGAATTTAGGCGAGATCTT 1699  
Db 961 GCGTCTTGGTTCTGGATGAACAAAAATTAAGCTATTTCAAGTGAATTTAGGCGAGATCTT 1020  
Qy 1700 TTGCCAAGCATGCCCTTAGAGAAGAACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTT 1759  
Db 1021 TTGCCAAGCATGCCCTTAGAGAAGAACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTT 1080  
Qy 1760 AATGCTTTGGATGCAAGCTTTTAAAGGTGCCGTGAAGTCCCTTCATGAATTTGATCTT 1819  
Db 1081 AATGCTTTGGATGCAAGCTTTTAAAGGTGCCGTGAAGTCCCTTCATGAATTTGATCTT 1140  
Qy 1820 CAGCGTTTTTTTATGGAGATGGAATACAAAAACTCTTTGCTATGGAATTTGGAAGTA 1879  
Db 1141 CAGCAGTTTTTTTATGGAGATGGAATACAAAAACTCTTTGCTATGGAATTTGGAAGTA 1200  
Qy 1880 TTATAAACAACCTGCAGAAAGAAATATTCAGTGTATGATCATGACAGAAGTATCTCTATAAC 1939  
Db 1201 TTATAAACAACCTGCAGAAAGAAATATTCAGTGTATGATCATGACAGAAGTATCTCTATAAC 1260  
Qy 1940 TGCATTTTCAGTTCAGATGTTTTACTGTTCTCTAATCTCGCTCGACATCTTATTTGAAGAGCA 1999  
Db 1261 TGCATTTTCAGTTCAGATGTTTTACTGTTCTCTAATCTCGCTCGACATCTTATTTGAAGAGCA 1320

Qy 2000 GAATGTTATCTCTGTCACTACTGAAACTCTGCTAGAACTTTTACCTGAGTACTTTGGACAG 2059  
Db 1321 GAATGTTATCTCTGTCACTACTGAACTCTGCTAGAACTTTTACCTGAGTACTTTGGACAG 1380  
Qy 2060 GAAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAATTTGGGAAAGATATATGCAAT 2119  
Db 1381 GAAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAATTTGGGAAAGATATATGCAAT 1440  
Qy 2120 AATATGTGACCTAAAGTATATCTGATCAGCAAAACCCACAATATGGACAGAAATTAAG 2179  
Db 1441 AATATGTGACCTAAAGTATATCTGATCAGCAAAACCCACAATATGGACAGAAATTAAG 1500  
Qy 2180 AATGCACTTCTTGAAGGTTTTCGATCTTTTTTGAAGATCTTTACCTGTATGCAAGGAAT 2239  
Db 1501 AATGCACTTCTTGAAGGTTTTCGATCTTTTTTGAAGATCTTTACCTGTATGCAAGGAAT 1560  
Qy 2240 GGAAGAAATCCGAAGACAGGTTGGGCAACAATTTGAAGTGGATCTCTGATTTGGGAGGCTGC 2299  
Db 1561 GGAAGAAATCCGAAGACAGGTTGGGCAACAATTTGAAGTGGATCTCTGATTTGGGAGGCTGC 1620  
Qy 2300 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTTCCAAGAGTGTGTGCTTG 2359  
Db 1621 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTTCCAAGAGTGTGTGCTTG 1680  
Qy 2360 TGATGAAGAACTCTTACTTGTGGCTTATAAGAAATGTCAAAAGCTGTGATGAGTGCAG 2419  
Db 1681 TGATGAAGAACTCTTACTTGTGGCTTATAAGAAATGTCAAAAGCTGTGATGAGTGCAG 1740  
Qy 2420 TACCAGTTTCATATCTAGTAGCAAGACAGTAGTACAATCTGTGTGGAATATGATTTGGAAAC 2479  
Db 1741 TACCAGTTTCATATCTAGTAGCAAGACAGTAGTACAATCTGTGTGGAATATGATTTGGAAAC 1800  
Qy 2480 AAGTCTCTACAGATATCTGAGGATCTTGTAAAGCATACATCTGCACTCTCTAGACCCCT 2539  
Db 1801 AAGTCTCTACAGATATCTGAGGATCTTGTAAAGCATACATCTGCACTCTCTAGACCCCT 1860  
Qy 2540 TGCTGGTCTTTCATGTAAGCTTTAAGCAGGCTGGGTGCTGTTTCAAGAATGCAATGATTTGT 2599  
Db 1861 TGCTGGTCTTTCATGTAAGCTTTAAGCAGGCTGGGTGCTGTTTCAAGAATGCAATGATTTGT 1920  
Qy 2600 GTCTTTTGAAGCACTTCAAGTAGAGGTAATGATGAGGAAATATCTTTTATGCTGTGGTGT 2659  
Db 1921 GTCTTTTGAAGCACTTCAAGTAGAGGTAATGATGAGGAAATATCTTTTATGCTGTGGTGT 1980  
Qy 2660 GGTGGCCAGGTTGCTGAGATGTGGCGAAGAAATGGAGTGTCTCTTATTTAGCCAGGT 2719  
Db 1981 GGTGGCCAGGTTGCTGAGATGTGGCGAAGAAATGGAGTGTCTCTTATTTAGCCAGGT 2040  
Qy 2720 GTTTTATTTACCAAGATGTTAAGTGCAGAGAAATGTATGATAAAGATATCATCATGCT 2779  
Db 2041 GTTTTATTTACCAAGATGTTAAGTGCAGAGAAATGTATGATAAAGATATCATCATGCT 2100  
Qy 2780 TCAGATTTGGTGCATTTTAAATGGAATCCCAATAAGTCTTGTGTTACTGTTAGAGGTA 2839  
Db 2101 TCAGATTTGGTGCATTTTAAATGGAATCCCAATAAGTCTTGTGTTACTGTTAGAGGTA 2160  
Qy 2840 TGAATTTGCCGAGGCTTTTAAAGACATATCTACAAAGACCAAGGATTTGATTTAAACA 2899  
Db 2161 TGAATTTGCCGAGGCTTTTAAAGACATATCTACAAAGACCAAGGATTTGATTTAAACA 2220  
Qy 2900 ATATAATACACTTAATAGAAGAAATGCTTCAGGTCTCATCTATATTTGGGTGAGCGTTA 2959  
Db 2221 ATATAATACACTTAATAGAAGAAATGCTTCAGGTCTCATCTATATTTGGGTGAGCGTTA 2280  
Qy 2960 TGTAACCTGGAGTGGGAAATGTGACCAAGAGAGGTCACAAATGAGAGAAATCATTCATT 3019  
Db 2281 TGTAACCTGGAGTGGGAAATGTGACCAAGAGAGGTCACAAATGAGAGAAATCATTCATT 2340  
Qy 3020 GCTTTGCAATTTGAAACCCATGCCACAGTGCATTTGCCAAAAATTTTACTCTGAGAGATGAAA 3079  
Db 2341 GCTTTGCAATTTGAAACCCATGCCACAGTGCATTTGCCAAAAATTTTACTCTGAGAGATGAAA 2400

3080 TAATGAACCTGGCTTAGAGATGTCTATAAACAAGTGGCCACATTTTAAAGAACCGAGTGT 3139  
Db TAATGAACCTGGCTTAGAGATGTCTATAAACAAGTGGCCACATTTTAAAGAACCGAGTGT 2460  
3140 ATCAGGCCATGGAGTTTATGAACATAAAGATGAATCACTGGAAGACTTCAATATATCTACTT 3199  
Db ATCAGGCCATGGAGTTTATGAACATAAAGATGAATCACTGGAAGACTTCAATATATCTACTT 2520  
3200 TTATCATTTATCTCAAAACCCAGCATAGCAGGCTGAACATATGCGAGAGAAAAGAGAGAAA 3259  
Db TTATCATTTATCTCAAAACCCAGCATAGCAGGCTGAACATATGCGAGAGAAAAGAGAGAAA 2580  
3260 ACAAGAAACAAGATGAAGCATTCGCCGCCACACCTCTCTGAAATCTGCTCCCTGCTTT 3319  
Db ACAAGAAACAAGATGAAGCATTCGCCGCCACACCTCTCTGAAATCTGCTCCCTGCTTT 2640  
3320 CAGCAAGATGATTAACCTTTCTCAACTGTATATCATATGATGATATCTCAGGACCGCTATT 3379  
Db CAGCAAGATGATTAACCTTTCTCAACTGTATATCATATGATGATATCTCAGGACCGCTATT 2700  
3380 TGAGCGGCAATAGACACAGATTTCTAATCTGTGGAACGAGGATGCTCCAAATGGCTTT 3439  
Db TGAGCGGCAATAGACACAGATTTCTAATCTGTGGAACGAGGATGCTCCAAATGGCTTT 2760  
3440 TCATATTTCTGGCATTTGGCTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCTCTGAAGA 3499  
Db TCATATTTCTGGCATTTGGCTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCTCTGAAGA 2820  
3500 AGAAGTAACTTTGATTTTATCATAGGCTTCAAGATTTGGAACTTCAGGCATGATAT 3559  
Db AGAAGTAACTTTGATTTTATCATAGGCTTCAAGATTTGGAACTTCAGGCATGATAT 2880  
3560 ACAATGCTTTTGGAAAACCTCAAGGAATTTCCCGATTTAGAGGCCAGAGGACATGAT 3619  
Db ACAATGCTTTTGGAAAACCTCAAGGAATTTCCCGATTTAGAGGCCAGAGGACATGAT 2940  
3620 AACCTGGATCTTCAGATGTTTTCACACAGTGAAGCGATTAAGAGAAAATCTTGTTTAAT 3679  
Db AACCTGGATCTTCAGATGTTTTCACACAGTGAAGCGATTAAGAGAAAATCTTGTTTAAT 3000  
3680 TGTAGCAACCATCAGGATCGAATCTATTAAAGATGATGATGATGATGATGATGATGATGAT 3738  
Db TGTAGCAACCATCAGGATCGAATCTATTAAAGATGATGATGATGATGATGATGATGATGAT 3059

RESULT 11  
LOCUS AK056441  
DEFINITION Homo sapiens cDNA FLJ31879 fis, clone NT2RP7002619, highly similar to Mus musculus ubiquitin-protein ligase E3-alpha (Ubr1) mRNA.  
ACCESSION AK056441  
VERSION AK056441.1 GI:16551844  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.  
1  
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakatsuki, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hoshi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotate, T., Kusanagi, J., Kanehori, K., Takahashi, Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mueshino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,

Shiohata, N., Sano, S., Moriya, S., Momiya, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, S., Itoh, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Negase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs Nat. Genet. 36 (1), 40-45 (2004)  
14702039  
2 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakatsuki, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T. NEDO human cDNA sequencing project Unpublished  
3 (bases 1 to 3059)  
Isogai, T., Otsuki, T. and Sugiyama, T. Direct Submission  
Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.  
Location/Qualifiers  
1. 3059  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP7002619"  
/cell\_line="NT2"  
/clone\_lib="NT2RP7"  
/note="cloning vector: pME18SPL3-mRNA from NT2 neuronal precursor cells after 5-weeks retinoic acid (RA) induction."

FEATURES  
source

Query Match 48.4%; Score 3055.8; DB 9; Length 3059;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 680 GTCAGTTTCCCTCAAGATGCGGACGAGGCGCTGGAGGTACTGAGAGGATGGAATCAG 739  
Db 1 GTCAGTTTCCCTCAAGATGCGGACGAGGCGCTGGAGGTACTGAGGAGTGAATCAG 60

## ORIGIN

Qy 740 CGCGAGTTACCCAGACCCCTTCAGCTCTGGCATCTTGGGGATCAGCAAGTTGATT 799  
Db 61 CGCGAGTTACCCAGACCCCTTCAGCTCTGGCATCTTGGGGATCAGCAAGTTGATT 120  
Qy 800 TTATCTGCTTCTTGTGATCATTTTGGCAATTTGGTGCAGAAATTTTACTTTGCTGAAT 859  
Db 121 TTATCTGCTTCTTGTGATCATTTTGGCAATTTGGTGCAGAAATTTTACTTTGCTGAAT 180

Qy 860 GGACCCAGACTTGGAAAGCAGGAGGAAGTGTACAAATGTCAATATTTCACTCCACTGGA 919  
Db |||||  
Qy 181 GGAACCCAGACTTGGAAAGCAGGAGGAAGTGTACAAATGTCAATATTTCACTCCACTGGA 240  
Db |||||  
Qy 920 ATGCTACTTATTTGGAGAGAGTCCAGATATTTGCTTAGAGAAAATGGAAGCAGCTGGAGC 979  
Db |||||  
Qy 241 ATGCTACTTATTTGGAGAGAGTCCAGATATTTGCTTAGAGAAAATGGAAGCAGCTGGAGC 300  
Db |||||  
Qy 980 ATTTTCAGCTTTGTTGGAGGGTTTTTCAAGTGGAGAGACCACTTATTTCTGGAGGATG 1039  
Db |||||  
Qy 301 ATTTTCAGCTTTGTTGGAGGGTTTTTCAAGTGGAGAGACCACTTATTTCTGGAGGATG 360  
Db |||||  
Qy 1040 TGCAATTTGATCCAAATGCTACTCTGTATGGACTGCTTCCAGCAGCTGTTCTATAAAA 1099  
Db |||||  
Qy 361 TGCATTTGATCCAAATGCTACTCTGTATGGACTGCTTCCAGCAGCTGTTCTATAAAA 420  
Db |||||  
Qy 1100 TCATCGTTTACAAGATGCATATCTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGA 1159  
Db |||||  
Qy 421 TCATCGTTTACAAGATGCATATCTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGA 480  
Db |||||  
Qy 1160 GGCAATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTGGAAGAGCAGTACTATAAA 1219  
Db |||||  
Qy 481 GGCAATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTGGAAGAGCAGTACTATAAA 540  
Db |||||  
Qy 1220 AGAGAAATTCAGCTGTCCGTTGAAATGAAGAGTAAATGTTCCAAAGCCAGGAAAATATTTCC 1279  
Db |||||  
Qy 541 AGAGAAATTCAGCTGTCCGTTGAAATGAAGAGTAAATGTTCCAAAGCCAGGAAAATATTTCC 600  
Db |||||  
Qy 1280 TTCAAGTATATAAATATGTCGTAGAAATGCACTATATGGGAAGAGGAAAAGAACTGCCCTCC 1339  
Db |||||  
Qy 601 TTCAAGTATATAAATATGTCGTAGAAATGCACTATATGGGAAGAGGAAAAGAACTGCCCTCC 660  
Db |||||  
Qy 1340 TGAATCCAGATAGGAGGAAAATGAAGAGTACTATTTGTGCTCTTTCAATGATGAACA 1399  
Db |||||  
Qy 661 TGAATCCAGATAGGAGGAAAATGAAGAGTACTATTTGTGCTCTTTCAATGATGAACA 720  
Db |||||  
Qy 1400 CCATTTATATGACACAGTCATATACAGCTTCAAAAGAGCTCTTGACTGTGAGCTCGAGA 1459  
Db |||||  
Qy 721 CCATTTATATGACACAGTCATATACAGCTTCAAAAGAGCTCTTGACTGTGAGCTCGAGA 780  
Db |||||  
Qy 1460 GGCCCGAGTGCATACCACTGCCATTTGACAAAGAGGGTCTGCGGCTGTTAAAGCGGAGC 1519  
Db |||||  
Qy 781 GGCCCGAGTGCATACCACTGCCATTTGACAAAGAGGGTCTGCGGCTGTTAAAGCGGAGC 840  
Db |||||  
Qy 1520 TTATGCTGCTTCCGAGGAGCAAGAGAGATATAAAGAGTCAATTCAGAAAATGTCTCA 1579  
Db |||||  
Qy 841 TTATGCTGCTTCCGAGGAGCAAGAGAGATATAAAGAGTCAATTCAGAAAATGTCTCA 900  
Db |||||  
Qy 1580 ACATCCACTTCAATGTAGAGTATTAACACTCAGAGATATATGCTCATCAGAAAATTTGCTTT 1639  
Db |||||  
Qy 901 ACATCCACTTCAATGTAGAGTATTAACACTCAGAGATATATGCTCATCAGAAAATTTGCTTT 960  
Db |||||  
Qy 1640 CGCTTCTGCTTCTGGATGAACAAATTTATGAGCTATTTCAAGTGAATTTAGGCGAGATCTT 1699  
Db |||||  
Qy 961 CGCTTCTGCTTCTGGATGAACAAATTTATGAGCTATTTCAAGTGAATTTAGGCGAGATCTT 1020  
Db |||||  
Qy 1700 TTGCCAAGCAGTCTTAGAGAGAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTT 1759  
Db |||||  
Qy 1021 TTGCCAAGCAGTCTTAGAGAGAACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTT 1080  
Db |||||  
Qy 1760 AATGCTTTGGAGTGCAGAGCTTTATAAGGTGCCGTAAAGTCCCTTCATGAATTTGATCTT 1819  
Db |||||  
Qy 1081 AATGCTTTGGAGTGCAGAGCTTTATAAGGTGCCGTAAAGTCCCTTCATGAATTTGATCTT 1140  
Db |||||  
Qy 1820 CAGCAGTTTTTTTATGGAGATGGATAACAAAACTCTTTGCTATGGAAATTTGGAAGTA 1879  
Db |||||  
Qy 1141 CAGCAGTTTTTTTATGGAGATGGATAACAAAACTCTTTGCTATGGAAATTTGGAAGTA 1200  
Db |||||  
Qy 1880 TTATAACACCTGCGAGAGAGATATATCAGTGAATCATGACAGAGTACTCTATAAC 1939  
Db |||||  
Qy 1201 TTATAACACCTGCGAGAGAGATATATCAGTGAATCATGACAGAGTACTCTATAAC 1260  
Db |||||  
Qy 1940 TGCACCTTTTCAGTATCAGATGTTTACTGTTCTCTACTCTCGCTCGACATCTTATTTGAAGAGCA 1999  
Db |||||

Db 1261 TGCACCTTTTCAGTATCAGATGTTTACTGTTCTCTACTCTCGCTCGACATCTTATTTGAAGAGCA 1320  
Qy |||||  
Qy 2000 GAATGTTATCTCTGTCTAATCTGAACTCTGCTAGAGTTTACCTCAGTACTTTGGACAG 2059  
Db |||||  
Qy 1321 GAATGTTATCTCTGTCTAATCTGAACTCTGCTAGAAAGTTTTACCTCAGTACTTTGGACAG 1380  
Db |||||  
Qy 2060 GAACAAATAAATCAACTTTCAAGGTTATAGCCAGGACAAAATTTGGGAAGAGTATATGCAGT 2119  
Db |||||  
Qy 1381 GAACAAATAAATCAACTTTCAAGGTTATAGCCAGGACAAAATTTGGGAAGAGTATATGCAGT 1440  
Db |||||  
Qy 2120 AATAATGACCTAAAGATPATATCTGTATCAGCAAAACCAACAATATGGAAGAGATTAAG 2179  
Db |||||  
Qy 1441 AATATGTGACTTAAAGTATATCTCTGTATAGCCAGGACAAAATTTGGGAAGAGTATTAAG 1500  
Db |||||  
Qy 2180 AATGCACTTCTTGAAGGTTTTCGATCTTTTGAAGATTTCTTACCTGTATGCAAGGAT 2239  
Db |||||  
Qy 1501 AATGCACTTCTTGAAGGTTTTCGATCTTTTGAAGATTTCTTACCTGTATGCAAGGAT 1560  
Db |||||  
Qy 2240 GGAAGAAATCCGAAGACAGGTTGGGCAACAATTTGAAGTGGATCTCTGATTTGGGAGGCTGC 2299  
Db |||||  
Qy 1561 GGAAGAAATCCGAAGACAGGTTGGGCAACAATTTGAAGTGGATCTCTGATTTGGGAGGCTGC 1620  
Db |||||  
Qy 2300 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTCAAGAGTGGTGTCTTG 2359  
Db |||||  
Qy 1621 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTCAAGAGTGGTGTCTTG 1680  
Db |||||  
Qy 2360 TGATGAAGAACTCTTACTTCTGCTTAAAGATGTCANAGCTGTGATGAGGTCGAG 2419  
Db |||||  
Qy 1681 TGATGAAGAACTCTTACTTCTGCTTAAAGATGTCANAGCTGTGATGAGGTCGAG 1740  
Db |||||  
Qy 2420 TACCAGTTTCATATCTAGTAGCAAGACAGTAGTAGTACAATCTGTTGGAATTTGGAAAC 2479  
Db |||||  
Qy 1741 TACCAGTTTCATATCTAGTAGCAAGACAGTAGTAGTACAATCTGTTGGAATTTGGAAAC 1800  
Db |||||  
Qy 2480 AAAGTCTTACAGAGTATCTGAGGATCTTTGAAGACATATCTGCTGCACTCTCTAGGACCT 2539  
Db |||||  
Qy 1801 AAAGTCTTACAGAGTATCTGAGGATCTTTGAAGACATATCTGCTGCACTCTCTAGGACCT 1860  
Db |||||  
Qy 2540 TGCTGCTTTCATGCTAGCTTTAAGCAGGCTGGGTCGTTTCAAGATGCAATGATTTGT 2599  
Db |||||  
Qy 1861 TGCTGCTTTCATGCTAGCTTTAAGCAGGCTGGGTCGTTTCAAGATGCAATGATTTGT 1920  
Db |||||  
Qy 2600 GTCTTTTGGAGACTTTCAGTGTAGAGTACTAGTGGAAATATCTTTTACTGCTGCTGCTT 2659  
Db |||||  
Qy 1921 GTCTTTTGGAGACTTTCAGTGTAGAGTACTAGTGGAAATATCTTTTACTGCTGCTGCTT 1980  
Db |||||  
Qy 2660 GGTTGCCCGAGGTTGCTGAGATGTCGCGAAGAAAATGGAATGGAATGCTCTTATTAAGCAGGT 2719  
Db |||||  
Qy 1981 GGTTGCCCGAGGTTGCTGAGATGTCGCGAAGAAAATGGAATGGAATGCTCTTATTAAGCAGGT 2040  
Db |||||  
Qy 2720 GTTTTATTAACAGATGTTAAGTGCAGAGAGAAATGCTATGATTAAGATATCATCATGCT 2779  
Db |||||  
Qy 2041 GTTTTATTAACAGATGTTAAGTGCAGAGAGAAATGATATGATTAAGATATCATCATGCT 2100  
Db |||||  
Qy 2780 TCAGATTTGGTGCATCTTTTAAATGGAATCCAAATAAGTCTTTGTTACTGCTGCTCAGAGGTA 2839  
Db |||||  
Qy 2101 TCAGATTTGGTGCATCTTTTAAATGGAATCCAAATAAGTCTTTGTTACTGCTGCTCAGAGGTA 2160  
Db |||||  
Qy 2840 TGAATCTTGGCGAGGCTTTTAAACAGACCATATCTCAAAAAGACCAAGATTTGATTAACA 2899  
Db |||||  
Qy 2161 TGAATCTTGGCGAGGCTTTTAAACAGACCATATCTCAAAAAGACCAAGATTTGATTAACA 2220  
Db |||||  
Qy 2900 ATATAATACACTAATAGAGAAATGCTTCAAGTCTCTCATATTTGGGTGAGCGTTA 2959  
Db |||||  
Qy 2221 ATATAATACACTAATAGAGAAATGCTTCAAGTCTCTCATATTTGGGTGAGCGTTA 2280  
Db |||||  
Qy 2960 TGTACCTGGAGTGGAAAATGTGACAAAAGAGAGGTCAATATGAGAGAAATCATTCATT 3019  
Db |||||  
Qy 2281 TGTACCTGGAGTGGAAAATGTGACAAAAGAGAGGTCAATATGAGAGAAATCATTCATT 2340  
Db |||||  
Qy 3020 GCTTTGCAATTTGAACCCCATGCGACACAGTGGCAATTTGCCAAAATTTTACTGAGATGAAAA 3079  
Db |||||

Db 2341 GCTTTGCTTGAACCCATGCCACACAGTGCCTTCCCAAAATTTACCTGAGATGAATA 2400  
QY 3080 TAATGAATCTGCTTTAGAGATGTCTATAAACAAGTGGCCACATTTAAGAAACCAAGTGT 3139  
Db 2401 TAATGAATCTGCTTTAGAGATGTCTATAAACAAGTGGCCACATTTAAGAAACCAAGTGT 2460  
QY 3140 ATCAGGCCATGAGTTTATGAATTAAGATGAATCACTGAAGACTTCAATATGTACTT 3199  
Db 2461 ATCAGGCCATGAGTTTATGAATTAAGATGAATCACTGAAGACTTCAATATGTACTT 2520  
QY 3200 TTATCATTTACTCCAAACCCAGCATAGCAAGGCTGAACATATGCAAGAAAGAGAGAAA 3259  
Db 2521 TTATCATTTACTCCAAACCCAGCATAGCAAGGCTGAACATATGCAAGAAAGAGAGAAA 2580  
QY 3260 ACAAGAAACAAAGATGAAGCAATTCGCCGCCACACACACTCTGTAATTTCTGCCCTGCTT 3319  
Db 2581 ACAAGAAACAAAGATGAAGCAATTCGCCGCCACACACACTCTGTAATTTCTGCCCTGCTT 2640  
QY 3320 CAGCAAGTGTATTAACCTTCTCACTGTGATATCATGATGTACATTTCTCAGGACCGTATT 3379  
Db 2641 CAGCAAGTGTATTAACCTTCTCACTGTGATATCATGATGTACATTTCTCAGGACCGTATT 2700  
QY 3380 TGAGCGGCAATAGACACAGATTTCTAATCTGTGGACCGAAGGATGCTCCAAATGGCTTT 3439  
Db 2701 TGAGCGGCAATAGACACAGATTTCTAATCTGTGGACCGAAGGATGCTCCAAATGGCTTT 2760  
QY 3440 TCATATTTCTGGATTTGCTTTTACAGAGAGAGCAACAGCTTCAAAAGCTTCTGGAAGA 3499  
Db 2761 TCATATTTCTGGATTTGCTTTTACAGAGAGAGCAACAGCTTCAAAAGCTTCTGGAAGA 2820  
QY 3500 AGAAGTACATTTGATTTTATCATAAGCTTCAAGATTTGGAAGTTCAGCCATGAATAT 3559  
Db 2821 AGAAGTACATTTGATTTTATCATAAGCTTCAAGATTTGGAAGTTCAGCCATGAATAT 2880  
QY 3560 ACAATGCTTTTGGAAAACTCAAGAGAAATCCCGAGTTAGAGGCCAAGAGGACATCAT 3619  
Db 2881 ACAATGCTTTTGGAAAACTCAAGAGAAATCCCGAGTTAGAGGCCAAGAGGACATCAT 2940  
QY 3620 AACGTGGATCTCAGATGTTTGACACAGTGAAGCGATTAAGAGAAAAATCTTGTTTAAT 3679  
Db 2941 AACGTGGATCTCAGATGTTTGACACAGTGAAGCGATTAAGAGAAAAATCTTGTTTAAT 3000  
QY 3680 TGTAGCACCATCAGATCGGATCTATTAAAGATCATGATGATGATGATGATGATGATGATGAT 3738  
Db 3001 TGTAGCACCATCAGATCGGATCTATTAAAGATCATGATGATGATGATGATGATGATGATGAT 3059

## RESULT 12

BD156870  
LOCUS BD156870 2550 bp DNA linear PAT 17-JAN-2003  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD156870

VERSION BD156870.1 GI:27862628

KEYWORDS JP 2002191363-A/11713.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2550)

Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A 11713 09-JUL-2002;

HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/11713

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU

PI SAITO,

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KEIICHI NAGAI, TETSUJI OTSUKI

PC

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10, C12P21/02, C12Q1/68, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers (117)..(2549).

FT CDS

Location/Qualifiers

1..2550

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

## FEATURES

source

## ORIGIN

Query Match 40.4%; Score 2546.8; DB 6; Length 2550;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1189 ATCATGAACCTGGAGAGCAGTACTATAAAGAGAAATTCACGCTGTCGCTTGAATGAAG 1248

Db 1 ATCATGAACCTGGAGAGCAGTACTATAAAGAGAAATTCACGCTGTCGCTTGAATGAAG 60

QY 1249 AGGTAATTTGTCACAGCCAGGAAATATTTCTTCAGTGAATAAATATGCTGAGAAATGA 1308

Db 61 AGGTAATTTGTCACAGCCAGGAAATATTTCTTCAGTGAATAAATATGCTGAGAAATGA 120

QY 1309 CTATATGGAAGAGAGAAAGAACTGCTCTCTGAATCTCAGATAGAGGAGAGAAATGAAA 1368

Db 121 CTATATGGAAGAGAGAAAGAACTGCTCTCTGAATCTCAGATAGAGGAGAGAAATGAAA 180

QY 1369 GATTAATTTGTCCTTTTCAATGATGAACCACTTATATGACCACTCATATACAGCC 1428

Db 181 GATTAATTTGTCCTTTTCAATGATGAACCACTTATATGACCACTCATATACAGCC 240

QY 1429 TACAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTCGATACCACTGCATTGACA 1488

Db 241 TACAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTCGATACCACTGCATTGACA 300

QY 1489 AAGAGGTGCTGCGGCTGTTAAAGCGGAGCTTATGCTGCTTGCAGAGAGAAAGGAAG 1548

Db 301 AAGAGGTGCTGCGGCTGTTAAAGCGGAGCTTATGCTGCTTGCAGAGAGAAAGGAAG 360

QY 1549 ATATAAGAGTCATTGAGAAATGCTCTCAACATCCACTTCATGATGAGATGATTAACCT 1608

Db 361 ATATAAGAGTCATTGAGAAATGCTCTCAACATCCACTTCATGATGAGATGATTAACCT 420

QY 1609 CAGAGATTTAGCTCATCAGAAATTTGCTTTGCGCTTCTGCTTCTGCTGCTGCTGCTGCT 1668

Db 421 CAGAGATTTAGCTCATCAGAAATTTGCTTTGCGCTTCTGCTTCTGCTGCTGCTGCT 480

QY 1669 TGAGCTATTCAAGTGACTTTTAGGCAGATCTTTTCCAGCAATGCTTTAGAGAGAACCTG 1728

Db 481 TGAGCTATTCAAGTGACTTTTAGGCAGATCTTTTCCAGCAATGCTTTAGAGAGAACCTG 540

QY 1729 ACTCGGAGATCCCTGCTCTATAAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAG 1788

Db 541 ACTCGGAGATCCCTGCTCTATAAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAG 600

QY 1789 GTGCCCGTAAGATCCTTCATGAATTTGCTTTCAGCAGTCTTTTATGAGATGGAATACA 1848

Db 601 GTGCCCGTAAGATCCTTCATGAATTTGCTTTCAGCAGTCTTTTATGAGATGGAATACA 660

QY 1849 AAAAATCTTTTGTCTATGGAATTTTGTGAAGTATTATAAACAATGAGAGAAATATATCA 1908

Db 661 AAAAATCTTTTGTCTATGGAATTTTGTGAAGTATTATAAACAATGAGAGAAATATATCA 720

QY 1909 GTGATGATCATGACAGAGATATCTATACTGCACTTTTCACTGATGTTTACTGTTTC 1968

Db 721 GTGATGATCATGACAGAGATATCTATACTGCACTTTTCACTGATGTTTACTGTTTC 780

QY 1969 CTACTCTGGCTCGACATCTTTATGAAGAGCAGAGATTTATCTCTGCTATTACTGAACTC 2028

Db 781 CTACTCTGGCTCGACATCTTTATGAAGAGCAGAGATTTATCTCTGCTATTACTGAACTC 840



Qy	2029	TGCTAGAAGT	TTTACCTGAGT	ATCTTGAGCA	AGGAACAA	TAAATTTCA	ACTTCCAGG	TTATTA	2088
Db	841	TGCTAGAAGT	TTTACCTGAGT	ATCTTGAGCA	AGGAACAA	TAAATTTCA	ACTTCCAGG	TTATTA	900
Qy	2089	GCCAGGCAAA	TTGGGAAGAGT	ATATCGAGT	TAATATGTG	ACCTTAAAGT	TATCTCGATCA	2148	
Db	901	GCCAGGCAAA	TTGGGAAGAGT	ATATCGAGT	TAATATGTG	ACCTTAAAGT	TATCTCGATCA	960	
Qy	2149	GCAACCCCA	ATAATGACGA	AAAGATTAA	GAATGCA	GTTCCTTGA	AGGTTTTTCGATCTT	2208	
Db	961	GCAACCCCA	ATAATGACGA	AAAGATTAA	GAATGCA	GTTCCTTGA	AGGTTTTTCGATCTT	1020	
Qy	2209	TTTTGAAGAT	CTTTACTGTAT	GCAGGAA	TGGAA	GAATTCGGA	AGACAGGTTGGGCAAC	2268	
Db	1021	TTTTGAAGAT	CTTTACTGTAT	GCAGGAA	TGGAA	GAATTCGGA	AGACAGGTTGGGCAAC	1080	
Qy	2269	ACATTGAAGT	GGATCTGTAT	TGGAGGCTGCC	ATTTGCTAT	TACAGATGCA	ATTTGAAGNATA	2328	
Db	1081	ACATTGAAGT	GGATCTGTAT	TGGAGGCTGCC	ATTTGCTAT	TACAGATGCA	ATTTGAAGNATA	1140	
Qy	2329	TTTTTACTAT	GTGTTCCA	AGAGTGGTGT	CTGTGTGAT	GAAAGAACTCTT	ACTTTGTGGCTTATA	2388	
Db	1141	TTTTTACTAT	GTGTTCCA	AGAGTGGTGT	CTGTGTGAT	GAAAGAACTCTT	ACTTTGTGGCTTATA	1200	
Qy	2389	AAGAATGTC	CAAAAGCTGT	ATGAGGTC	AGTACCA	AGTTTTCATATCT	TAGTAGCAAGCAG	2448	
Db	1201	AAGAATGTC	CAAAAGCTGT	ATGAGGTC	AGTACCA	AGTTTTCATATCT	TAGTAGCAAGCAG	1260	
Qy	2449	TAGTACAAT	CGTGGGACA	TAGTTTGGAA	ACAAAGTCT	CACAGATATCT	GAGGATCTTG	2508	
Db	1261	TAGTACAAT	CGTGGGACA	TAGTTTGGAA	ACAAAGTCT	CACAGATATCT	GAGGATCTTG	1320	
Qy	2509	TAAGCATAC	ATCTGCCA	CTCTTAGG	ACCCTGCTG	CTTCAATG	TACGTTTAAAGCAGGC	2568	
Db	1321	TAAGCATAC	ATCTGCCA	CTCTTAGG	ACCCTGCTG	CTTCAATG	TACGTTTAAAGCAGGC	1380	
Qy	2569	TGGGTGCTG	TTTCAAGACT	GCATGAA	TTTGTGTCTTTT	TGAGGA	CTTTTCAAGTAGAGGTAC	2628	
Db	1381	TGGGTGCTG	TTTCAAGACT	GCATGAA	TTTGTGTCTTTT	TGAGGA	CTTTTCAAGTAGAGGTAC	1440	
Qy	2629	TAGTGGAA	TATCTTTTA	CGTTGTCTG	TGTTGGT	TGCCCAGG	TGTTGCTGAGATGNGGC	2688	
Db	1441	TAGTGGAA	TATCTTTTA	CGTTGTCTG	TGTTGGT	TGCCCAGG	TGTTGCTGAGATGNGGC	1500	
Qy	2689	GAAGAA	TGGA	CTCTCTTA	TTAGCCAG	TGTTTTAT	TACCAAGATGTTAAAGTGCAAG	2748	
Db	1501	GAAGAA	TGGA	CTCTCTTA	TTAGCCAG	TGTTTTAT	TACCAAGATGTTAAAGTGCAAG	1560	
Qy	2749	AAGAAATGT	ATGA	TAAAGATAT	CA	TATGCTT	CAGATG	TG	2808
Db	1561	AAGAAATGT	ATGA	TAAAGATAT	CA	TATGCTT	CAGATG	TG	1620
Qy	2809	ATAAGT	CTTGT	TTACTGG	TA	CTCAGAGG	TATGAA	CTTGGCGAGG	2868
Db	1621	ATAAGT	CTTGT	TTACTGG	TA	CTCAGAGG	TATGAA	CTTGGCGAGG	1680
Qy	2869	TATCTCA	AAAGACCA	GGATTTG	ATAA	CAATATAT	TAC	TAATAG	2928
Db	1681	TATCTCA	AAAGACCA	GGATTTG	ATAA	CAATATAT	TAC	TAATAG	1740
Qy	2929	AGTCTCT	CATCTAT	ATTTGT	G	G	G	G	2988
Db	1741	AGTCTCT	CATCTAT	ATTTGT	G	G	G	G	1800
Qy	2989	AAGAGG	TCA	CAATGAG	AAATCA	TTGCTTTG	CA	ATTTGA	3048
Db	1801	AAGAGG	TCA	CAATGAG	AAATCA	TTGCTTTG	CA	ATTTGA	1860
Qy	3049	CCATTTGCC	AAAAATTT	ACCTGAG	AAATG	AAAAATAT	TGAA	CTGGCT	3108
Db	1861	CCATTTGCC	AAAAATTT	ACCTGAG	AAATG	AAAAATAT	TGAA	CTGGCT	1920
Qy	3109	ACAAAGT	GGCCACAT	TTTAA	GA	AAAC	CCAGG	TATCAGG	3168

Db	1921	ACAAAGTGGGCCAATTTAAAGAAACCGAGGTGATCAGGCCATGGAGTTTATGAACATAAAG	1980
Qy	3169	ATGAATCACTGAAGAGCTTCAAATGATGACTTTTATCATTTACTCTCAAAACCCAGCATAGCA	3228
Db	1981	ATGATCACTGAAGAGCTTCAAATGATGACTTTTATCATTTACTCTCAAAACCCAGCATAGCA	2040
Qy	3229	AGGCTGAACATATGCGAAGAAAGGAGAGAAAACAAGAAAACAAAGATGAAGCAATTCGCCG	3288
Db	2041	AGGCTGAACATATGCGAAGAAAGGAGAGAAAACAAGAAAACAAAGATGAAGCAATTCGCCG	2100
Qy	3289	CACCAACCACTCTGGAATTCCTGCCCTGCTTTTCAGCAAGAGTGATTAACCTTCTCAACTGTG	3348
Db	2101	CACCAACCACTCTGGAATTCCTGCCCTGCTTTTCAGCAAGAGTGATTAACCTTCTCAACTGTG	2160
Qy	3349	ATATCATGATGTACATTTCTCAGGACCGTATTTTCAGCGGGCAATAGACACAGATTTCTAACT	3408
Db	2161	ATATCATGATGTACATTTCTCAGGACCGTATTTTCAGCGGGCAATAGACACAGATTTCTAACT	2220
Qy	3409	TGTGGACCGAAGGGATGCTCCAAATGCTTTTCATATTTCTGGCAATTCGGTTTACTAGAAG	3468
Db	2221	TGTGGACCGAAGGGATGCTCCAAATGCTTTTCATATTTCTGGCAATTCGGTTTACTAGAAG	2280
Qy	3469	AGAAGCAACAGCTTCAAAAAGCTCTCGAAGAGAGTAAACATTTTGACTTTTATCATTAAG	3528
Db	2281	AGAAGCAACAGCTTCAAAAAGCTCTCGAAGAGAGTAAACATTTTGACTTTTATCATTAAG	2340
Qy	3529	CTTCAAGATTTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAA	3588
Db	2341	CTTCAAGATTTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAA	2400
Qy	3589	TTCCCCAGTTAGAAGGCCAGAGGACATGATAACGTCGTGATCTTCAGATGTTTGACACAG	3648
Db	2401	TTCCCCAGTTAGAAGGCCAGAGGACATGATAACGTCGTGATCTTCAGATGTTTGACACAG	2460
Qy	3649	TGAAGCGATTAGAGAAAAATCTTGTGTTAATTTGTAGCAACCAATCAGGATCGGAATCTA	3708
Db	2461	TGAAGCGTTTAAAGAAAAATCTTGTGTTAATTTGTAGCAACCAATCAGGATCGGAATCTA	2520
Qy	3709	TTAAGATGATGAGATTTACTCATGATAAAG	3738
Db	2521	TTAAGATGATGAGATTTACTCATGATAAAG	2550
RESULT	13		
AX877826			
LOCUS	AX877826	2550 bp	DNA
DEFINITION	Sequence 12731 from Patent EP1074617.		linear
ACCESSION	AX877826		
VERSION	AX877826.1	GI:40032562	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Oca, T., Isoqai, T., Nishikawa, T., Hayashi, K., Saico, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Oeueki, T.		
TITLE	Primers for synthesising full-length cDNA and their use		
JOURNAL	Patent: EP 1074617-A 12731 07-FEB-2001.		
FEATURES	Research Association for Biotechnology (JP)		
source	Location/Qualifiers		
	1..2550		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	117..>2549		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CA892293.1"		
	/db_xref="GI:40032563"		
	/translation="MTIWEEEKSLPPELQITRENNRYCYVLFNDEHSHYDHVITYSLROR		
	ALPCELABAQLHTTADIKGERRAVTGAAYAACAKEDIKSHSNVSOQLPHVFLVLS		
CDS			

[illegible]

## ORIGIN

Query Match 40.4%: Score 2546.8: DB 6: Length 2550:

Query Match	40.44	Score 2340.8
Best Local Similarity	99.94	Pred No. 0

100% local similarity 55.5%; read no: 0;  
 Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1189	ATCATGAACCTGGAGAGCAGGTACTATATAAAGAGAAATTCACGCTGTCCTTGAATGAAG	1248
Db	1	ATCATGAACCTGGAGAGCAGGTACTATAAAGAGAAATTCACGCTGTCCTTGAATGAAG	60
Qy	1249	AGGTAAATGTCCAAAGCCAGGAAAATATTTCTTCAGTGATAAAATATGTCGTAGAATGA	1308
Db	61	AGGTAAATGTCCAAAGCCAGGAAAATATTTCTTCAGTGATAAAATATGCGGTAGAATGA	120
Qy	1309	CTATATGGGAAGAGGAAAAAGAACTGCTCTCGAATCTCCAGATAAGGGAGAAAAATGAAA	1368
Db	121	CTATATGGGAAGAGGAAAAAGAACTGCTCTCGAATCCAGATAGGGAGAAAAATGAAA	180
Qy	1369	GATACTAATGTGTCCTTTTCAATGATGAACAACCAATTCATATGACCAAGTCATATACAGCC	1428
Db	181	GATACTAATGTGTCCTTTTCAATGATGAACAACCAATTCATATGACCAAGTCATATACAGCC	240
Qy	1429	TACAAAGAGCTCTTGACTGTGAGCTGCGACAGGCCAGTTGCAATCCACTGCCATTGACA	1488
Db	241	TACAAAGAGCTCTTGACTGTGAGCTGCGACAGGCCAGTTGCAATCCACTGCCATTGACA	300
Qy	1489	AAGAGGGTCTGCGGGCTGTTTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAAAGAAAG	1548
Db	301	AAGAGGGTCTGCGGGCTGTTTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAAAGAAAG	360
Qy	1549	ATATAAGAGTCATTACAGAAAATGTCTCTCAACATCCACTTCATGTAGAGTATTACACT	1608
Db	361	ATATAAGAGTCATTACAGAAAATGTCTCTCAACATCCACTTCATGTAGAGTATTACACT	420
Qy	1609	CAGAGATTATGGCTCATCAGAAATTTGCTTTGGCTCTTGCTGCTGATGAACAAAAATTA	1668
Db	421	CAGAGATTATGGCTCATCAGAAATTTGCTTTGGCTCTTGCTGCTGATGAACAAAAATTA	480
Qy	1669	TGAGCTATTCCAAGTGACTTTTAGCCAGATCTTTTGGCCAAGCATGCCCTTAGAGAGAACCTG	1728
Db	481	TGAGCTATTCCAAGTGACTTTTAGGGAGATCTTTTGGCCAAGCATGCCCTTAGAGAGAACCTG	540
Qy	1729	ACTCGGAGAACTCCTGTCTCATTAAGCAGGTTAATGCTTTGGGATGCAAAAGCTTTATAAG	1788
Db	541	ACTCGGAGAACTCCTGTCTCATTAAGCAGGTTAATGCTTTGGGATGCAAAAGCTTTATAAG	600
Qy	1789	GTGCGCGTAAGATCCTTCATGAANTGATCTTACGACAGTTTTTTTTTATGGAGATGGAATACA	1848
Db	601	GTGCGCGTAAGATCCTTCATGAANTGATCTTACGACAGTTTTTTTTTATGGAGATGGAATACA	660
Qy	1849	AAAAACTCTTTGCTATGGAAATTTGTGAAGTATTATAAACAACCTGCAGAAAGATATATCA	1908
Db	661	AAAAACTCTTTGCTATGGAAATTTGTGAAGTATTATAAACAACCTGCAGAAAGATATATCA	720
Qy	1909	GTGATGATCATGACAGAAAGTATCTCTATACTGCACCTTTTCAGTTCAGATGTTTACTGTTTC	1968
Db	721	GTGATGATCATGACAGAAAGTATCTCTATACTGCACCTTTTCAGTTCAGATGTTTACTGTTTC	780
Qy	1969	CTACTCTGGCTCGACATCTTATTGAAGAGCAGAAATGTTATCTCTGTCTATTACTGAAACTC	2028

```
QY 3109 ACAAGTGGCCACATTTTAAGAAACACAGGTGTATCAGGCCATGGAGTTTATGAACCTAAAG 3168
Db 1921 ACAAGTGGCCACATTTTAAGAAACACAGGTGTATCAGGCCATGGAGTTTATGAACCTAAAG 1980
QY 3169 ATGAATCACTGAAGACTTCAATATGTACTTTTATCATTACTCCAAAACCCAGCATAGCA 3228
Db 1981 ATGAATCACTGAAGACTTCAATATGTACTTTTATCATTACTCCAAAACCCAGCATAGCA 2040
QY 3229 AGGCTGAACATATGCAAGAAAGAGAGAAACAAAGAAACAAAGAAACAAAGATGAGCATGGCCG 3288
Db 2041 AGGCTGAACATATGCAAGAAAGAGAGAAACAAAGAAACAAAGATGAGCATGGCCG 2100
QY 3289 CACCAACCACTCTCTAAATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3348
Db 2101 CACCAACCACTCTCTAAATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
QY 3349 ATATCATGATGTACATTTCTCAGGACCGGTATTTGAGCGGGCAATAGACACAGATTTCAACT 3408
Db 2161 ATATCATGATGTACATTTCTCAGGACCGGTATTTGAGCGGGCAATAGACACAGATTTCAACT 2220
QY 3409 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGTCATGTTGGCATGGTTTACTAGAAG 3468
Db 2221 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGTCATGTTGGCATGGTTTACTAGAAG 2280
QY 3469 AGACACACACGCTTCAAAAGCTCTGAGAGAGAGTAACTTTGACATTTTATCATTAAGG 3528
Db 2281 AGACACACACGCTTCAAAAGCTCTGAGAGAGAGTAACTTTGACATTTTATCATTAAGG 2340
QY 3529 CTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAATCTCAAAGGAA 3588
Db 2341 CTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAATCTCAAAGGAA 2400
QY 3589 TTCCCGAGTTAGAGCCGAGAGACATGATTAACGTGGATACCTTCAGATGTTTGCACAG 3648
Db 2401 TTCCCGAGTTAGAGCCGAGAGACATGATTAACGTGGATACCTTCAGATGTTTGCACAG 2460
QY 3649 TGAACGGATTAGAGAAAAATCTCTTAAATGTTAGCAACCAACATCAGGATCGGAATCTA 3708
Db 2461 TGAACGGATTAGAGAAAAATCTCTTAAATGTTAGCAACCAACATCAGGATCGGAATCTA 2520
QY 3709 TTAAGATGATGAGATTAATCTCATGATAAAG 3738
Db 2521 TTAAGATGATGAGATTAATCTCATGATAAAG 2550

RESULT 14
AK027803
LOCUS
DEFINITION Homo sapiens cDNA FLJ14897 fig. clone PLACE1004743, weakly similar
to PROBABLE N-END-RECOGNIZING PROTEIN.
ACCESSION AK027803
VERSION AK027803.1 GI:14042751
KEYWORDS oligo capping; fig (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, T., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
```

```
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Teraehima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, P., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
2
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y.,
Matsunawa, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y. and Sasaki, N.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2550)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
1. .2550
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004743"
/tissue_type="placenta"
/clone_lib="PLACE1"
/notes="cloning vector: pME18SFL3"
117. .>2550
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB55380.1"
/db_xref="GI:14042752"
/translaton="MTIWEKEKLPPELQIRKERNRYCYVLFNDEHHSYDHYVLSQRL
IQMKKNILMFQWCACDELLLVAYKCHKAVMRCSTSPISSTKTVQSCGHSLET
KSVKSDVLSIHLPSRLTLAGLVRLSGLGSRHLHFVSPDFQVEVLVYPLRLCL
LVQYVAEMWRRNGLSISQVFFYQDVCKREMYDKDIIMLQIGASLMDPNKELLV
LORVLAELAEFKNTISTKQDVLIKOYNTLIERMLQVLIYICGERVPGVGNVKEEVTM
RELIHLICIEPMFSAIXNLIPENNETGLENVINKVATFKPGVSGHYVELKDES
LKDPMYFHYKSTQHSKAEHMQCKRKNQENKORALPPPPPPFCFPKPGVSGHYVELKDES
IMWILTVFERAIDTDSNLWTEGMLQMAFHILALGLLEKEQLOKQAPBEBSVTFDYH
KASRLGSSAMNIQMLEKLGIPQLEGQKDMITWILQMFDTVKELREKSLIVATTS
SESINKDEIITHDK"
```

## FEATURES

source

## CDS

## ORIGIN

Query Match 40.4%; Score 2546.8; DB 9; Length 2550;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;





Db 2161 ATATCATGATGTATCATCTTCAGGACCGTATTTGAGCGGCAATAGACACAGATTCCTAACT 2220  
QY 3409 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGGCATTTGGGTTTACTAGAAG 3468  
Db 2221 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGGCATTTGGGTTTACTAGAAG 2280  
QY 3469 AGAAGCAACAGCTTCAAAGAGCTCTGGAAGAGAGTAACATTTTCGACTTTTATCATAAAGG 3528  
Db 2281 AGAAGCAACAGCTTCAAAGAGCTCTGGAAGAGAGTAACATTTTCGACTTTTATCATAAAGG 2340  
QY 3529 CTTCAAGATTTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAATCTCAAAGGAA 3588  
Db 2341 CTTCAAGATTTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAATCTCAAAGGAA 2400  
QY 3589 TTCCCCAGTTAGAAGCCGAGAAGGACATGATTAACCTGGATATCTTCAGATGTTTGAACAG 3648  
Db 2401 TTCCCCAGTTAGAAGCCGAGAAGGACATGATTAACCTGGATATCTTCAGATGTTTGAACAG 2460  
QY 3649 TGAAGCGATTAAAGAGAAAAATCTGTTTAAATTTGTAGCAACACATCAGGATCGGAATCTA 3708  
Db 2461 TGAAGCGATTAAAGAGAAAAATCTGTTTAAATTTGTAGCAACACATCAGGATCGGAATCTA 2520  
QY 3709 TTAAGAATGATGAGATTAATCTCATGATAAAG 3738  
Db 2521 TTAAGAATGATGAGATTAATCTCATGATAAAG 2550

RESULT 15  
HSM808548 4576 bp mRNA linear PRI 30-AUG-2003  
LOCUS Homo sapiens mRNA; cDNA DKFZp686M07224 (from clone DKFZp686M07224).  
DEFINITION BX648400  
ACCESSION BX648400.1 GI:34367562  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE Bloembergen, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
Fob, G., Han, M. and Wiemann, S.  
AUTHORS The German Human cDNA Consortium  
CONSTRM Direct Submission  
TITLE Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
JOURNAL Neuherberg, GERMANY  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFZp686M07224) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de  
information about the clone and the sequencing project is available  
at http://mips.gsf.de/proj/cDNA/.

FEATURES  
source  
1..4576  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686M07224"  
/tissue\_type="human small intestine"  
/clone\_lib="686 (synonym: hicc3). Vector pSport1\_Sfi; host  
DH10B; sites SfiIA + SfiIB"  
/dev\_stage="adult"  
polyA\_signal 4528..4533  
polyA\_site 4549

Query Match 39.4%; Score 2487.8; DB 9; Length 4576;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3818 GAAAAAATTTCATTTGAAAACTCATAACTCATGTATGACAAATACATACAGAAATGCCTGGAA 3877  
Db 1 GAAAAAATTTCATTTGAAAACTCATAACTCATGTATGACAAATACATACAGAAATGCCTGGAA 60  
QY 3878 AGAAGATTTCCATTTATGAGGAGAGAGACACCCAGCAGTCACTGAGTCACTCTAGAAATGC 3937  
Db 61 AGAAGATTTCCATTTATGAGGAGAGAGAGACACCCAGCAGTCACTGAGTCACTCTAGAAATGC 120  
QY 3938 TTTGGGTCCTTAAACCGGGTCCATCTGTTATCTGAAAAAGAGGTGCTGAGTGCATCCTTTG 3997  
Db 121 TTTGGGTCCTTAAACCGGGTCCATCTGTTATCTGAAAAAGAGGTGCTGAGTGCATCCTTTG 180  
QY 3998 CCAAGAGAGACAGAGAGGTGAAAAATAGAAAAATAATGCCATGTTATTCGGCTGTGTCCA 4057  
Db 181 CCAAGAGAGACAGAGAGGTGAAAAATAGAAAAATAATGCCATGTTATTCGGCTGTGTCCA 240  
QY 4058 GAAATCTACTGCTTTAAACCCAGACACAGGGGAAAAACCCATAGAACTCTCAGGAGAGGCCCT 4117  
Db 241 GAAATCTACTGCTTTAAACCCAGACACAGGGGAAAAACCCATAGAACTCTCAGGAGAGGCCCT 300  
QY 4118 AGACCCACTTTTCATGATCCAGACTTTGGCATATGGAATTTATACAGGAAGCTGTGTCA 4177  
Db 301 AGACCCACTTTTCATGATCCAGACTTTGGCATATGGAATTTATACAGGAAGCTGTGTCA 360  
QY 4178 TGTAAATGCACGAGTGTGTCGCGAGAAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCA 4237  
Db 361 TGTAAATGCACGAGTGTGTCGCGAGAAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCA 420  
QY 4238 GCGCAATTCATGTTGACCTTTTGAATTTGAAAGTGGAGAAATATCTTTGCCCTCTTTGCA 4297  
Db 421 GCGCAATTCATGTTGACCTTTTGAATTTGAAAGTGGAGAAATATCTTTGCCCTCTTTGCA 480  
QY 4298 ATCTCTGTCGAATGATGTCGATCCCATATTTCTTTTGGCACTCAAAGATAACAGCTGA 4357  
Db 481 ATCTCTGTCGAATGATGTCGATCCCATATTTCTTTTGGCACTCAAAGATAACAGCTGA 540  
QY 4358 GAATGAGAGATGCTCTGCTCAACTTTTGAACCTCGCAGCGTGGATACAGACTGTTTGGC 4417  
Db 541 GAATGAGAGATGCTCTGCTCAACTTTTGAACCTCGCAGCGTGGATACAGACTGTTTGGC 600  
QY 4418 CAGAAATTCAGGTTTAAATATAAGACATGCTTAAAGGAGAAAAACCAATTCCTATTTCTT 4477  
Db 601 CAGAAATTCAGGTTTAAATATAAGACATGCTTAAAGGAGAAAAACCAATTCCTATTTCTT 660  
QY 4478 TAATCAGGAATGGAGATTCCTTTGGAGTTCATTCATCTCTGAGTTTGGGTTGA 4537  
Db 661 TAATCAGGAATGGAGATTCCTTTGGAGTTCATTCATCTCTGAGTTTGGGTTGA 720  
QY 4538 GTCTTCGATTAATAATTCATAATAGCATCAAGGAAATGTTATTTCTTTGGCAACAACAT 4597  
Db 721 GTCTTCGATTAATAATTCATAATAGCATCAAGGAAATGTTATTTCTTTGGCAACAACAT 780  
QY 4598 TTATAGAAATTTGAAAGTGGCACTCTGATGAAAGGATCCTCGAGTCCCATCTGAC 4657  
Db 781 TTATAGAAATTTGAAAGTGGCACTCTGATGAAAGGATCCTCGAGTCCCATCTGAC 840  
QY 4658 CTGAGACACCTGCGCTTTCATCTATCCAGGCAATTTGAAAAATCTATTTGGAGATGAGGAA 4717  
Db 841 CTGAGACACCTGCGCTTTCATCTATCCAGGCAATTTGAAAAATCTATTTGGAGATGAGGAA 900  
QY 4718 ACCTCTGTTGGAGCCTTTCAAAATAGGCAATTAATGCTGAAAGCATTAATGTCAGTT 4777  
Db 901 ACCTCTGTTGGAGCCTTTCAAAATAGGCAATTAATGCTGAAAGCATTAATGTCAGTT 960  
QY 4778 TGCAGTTGCACAGAGGATTAATCTGCTCTCAGGTCTCTGATACAGAAACATCTGGTTCGT 4837  
Db 961 TGCAGTTGCACAGAGGATTAATCTGCTCTCAGGTCTCTGATACAGAAACATCTGGTTCGT 1020  
QY 4838 TCTATCAGTTGTTCTTCTTAACATAAAATCAGAGATACACCATGCTCTCTATAGA 4897  
Db 1021 TCTATCAGTTGTTCTTCTTAACATAAAATCAGAGATACACCATGCTCTCTATAGA 1080

QY 4898 TCTGTTTCATGTTTGGTGGGTGCTGTGTAGCATTCCTATCCCATCTCTGTATTTGGATGACCC 4957  
DB 1081 TCTGTTTCATGTTTGGTGGGTGCTGTGTAGCATTCCTATCCCATCTCTGTATTTGGATGACCC 1140  
QY 4958 TGTGATCTGCAGCCCTTCTTCAGTTAGTTCTTCCTATACCACTTTATCTCTCCATTT 5017  
DB 1141 TGTGATCTGCAGCCCTTCTTCAGTTAGTTCTTCCTATACCACTTTATCTCTCCATTT 1200  
QY 5018 GATCACCATTGGCACAATGCTTCAGATACTTACTTACAGTAGACAGAGGCTTACCCCTTGC 5077  
DB 1201 GATCACCATTGGCACAATGCTTCAGATACTTACTTACAGTAGACAGAGGCTTACCCCTTGC 1260  
QY 5078 TCAGGTTCAAGAGACAGTAGAGGCTCATTCGCGCATCTCTTCTTTCGAGAAATTC 5137  
DB 1261 TCAGGTTCAAGAGACAGTAGAGGCTCATTCGCGCATCTCTTCTTTCGAGAAATTC 1320  
QY 5138 TCAATATACAGTGGCTCATTTGGGTGATATTCCTGGCTGGTATTTGTGGGTCTCACT 5197  
DB 1321 TCAATATACAGTGGCTCATTTGGGTGATATTCCTGGCTGGTATTTGTGGGTCTCACT 1380  
QY 5198 GAAGAATGGCATCACCCCTTATCTTCGTGTGTGCTGCAATTTTTCACATATTACTTGG 5257  
DB 1381 GAAGAATGGCATCACCCCTTATCTTCGTGTGTGCTGCAATTTTTCACATATTACTTGG 1440  
QY 5258 GGTACTCCGCTGAGGAATGCTATACCAATTTCTGCAAGAGAGATACAGTGCATCTG 5317  
DB 1441 GGTAACTCCGCTGAGGAATGCTATACCAATTTCTGCAAGAGAGATACAGTGCATCTG 1500  
QY 5318 TAGCTATCTATCTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGATACTGT 5377  
DB 1501 TAGCTATCTATCTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGATACTGT 1560  
QY 5378 AAGGCCCTTGTCCAGAGGTGGTGTGAGATCCTGCTTACTTAAACTGTTTGAAGCAAAA 5437  
DB 1561 AAGGCCCTTGTCCAGAGGTGGTGTGAGATCCTGCTTACTTAAACTGTTTGAAGCAAAA 1620  
QY 5438 AAAACCGTGTGAGTACCTTAGAAAAAGAAATAGTTGTAGAGCTTCTGTAGACTA 5497  
DB 1621 AAAACCGTGTGAGTACCTTAGAAAAAGAAATAGTTGTAGAGCTTCTGTAGACTA 1680  
QY 5498 TAGTGCCTCTGGAATCAAGCTTCTCATTTCAAGTGCCACCGTCTGAGATGATGAGG 5557  
DB 1681 TAGTGCCTCTGGAATCAAGCTTCTCATTTCAAGTGCCACCGTCTGAGATGATGAGG 1740  
QY 5558 AAAGCATCCTGCTCTGCTCTTCTGTGGGCTATATATGTTCTCAGAACATTTGCTG 5617  
DB 1741 AAAGCATCCTGCTCTGCTCTTCTGTGGGCTATATATGTTCTCAGAACATTTGCTG 1800  
QY 5618 CAGGAAATTTGAAACGGGGAAGAGGTTGAGCTTGCAATTTTTCACGCACTTCACTGTGG 5677  
DB 1801 CAGGAAATTTGAAACGGGGAAGAGGTTGAGCTTGCAATTTTTCACGCACTTCACTGTGG 1860  
QY 5678 AGCCGGGTCTGCAATTTCTTAAATCAGAGATCCGAGTGGTCTGTTGAAGGTAA 5737  
DB 1861 AGCCGGGTCTGCAATTTCTTAAATCAGAGATCCGAGTGGTCTGTTGAAGGTAA 1920  
QY 5738 AGCCAGAGCTGTGCTATCCAGCTCTTACTTGTGATGAATGAGAAACAGACCCCTGG 5797  
DB 1921 AGCCAGAGCTGTGCTATCCAGCTCTTACTTGTGATGAATGAGAAACAGACCCCTGG 1980  
QY 5798 CTTGAAGGGGCAACCCCTTTCATTTATCTCGTAGCGGTATCGGAAGCTCCATTTGGT 5857  
DB 1981 CTTGAAGGGGCAACCCCTTTCATTTATCTCGTAGCGGTATCGGAAGCTCCATTTGGT 2040  
QY 5858 CTGGCAACACACTGCATTTATAGAGAGATTTGCTAGGAGCCAGAGACTAATCAGATGTT 5917  
DB 2041 CTGGCAACACACTGCATTTATAGAGAGATTTGCTAGGAGCCAGAGACTAATCAGATGTT 2100  
QY 5918 ATTTGATTTCAACTGCGAGTTACTGTGAGCTCCAACTCTGCTCAAGACAATCACAATG 5977  
DB 2101 ATTTGATTTCAACTGCGAGTTACTGTGAGCTCCAACTCTGCTCAAGACAATCACAATG 2160  
QY 5978 ACACAGTAGTAAAGGCTGATTTCAAAATTTATGGAATACTTTCTGAGGGCTGGGAAGTAT 6037

DB 2161 ACACAGTAGTAAAGGCTGATTTCAAAATTTATGGAATACTTTCTGAGGGCTGGGAAGTAT 2220  
QY 6038 TGGAGGGTCTTTTGTCTCCATGTCCAGGTTCACTTACATCAATAAATAATTTCTTAATGA 6097  
DB 2221 TGGAGGGTCTTTTGTCTCCATGTCCAGGTTCACTTACATCAATAAATAATTTCTTAATGA 2280  
QY 6098 GTATTGCTTTCAATTTAGCAAAATATGCTTCAAGGAAAAAGGACATAGATCAATCTGT 6157  
DB 2281 GTATTGCTTTCAATTTAGCAAAATATGCTTCAAGGAAAAAGGACATAGATCAATCTGT 2340  
QY 6158 TTTATGCTAGTATTTCCAGGAATTTATTCCTTCAATTTGCTCTCATTTTCAATTTA 6217  
DB 2341 TTTATGCTAGTATTTCCAGGAATTTATTCCTTCAATTTGCTCTCATTTTCAATTTA 2400  
QY 6218 TTTCAATCCACTTGGTAGATGAAGTCACTCAAAACAGTTGTAGACATTTTATGTTGGT 6277  
DB 2401 TTTCAATCCACTTGGTAGATGAAGTCACTCAAAACAGTTGTAGACATTTTATGTTGGT 2460  
QY 6278 AACTCTTCTGCAATTTTGTATTTGGTGT 6308  
DB 2461 AACTCTTCTGCAATTTTGTATTTGGTGT 2491

Search completed: July 9, 2005, 02:53:33  
Job time : 25881 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 18:33:52 ; Search time 2893 Seconds  
(without alignments)  
12907.614 Million cell updates/sec

Title: US-10-758-672A-1  
Perfect score: 6308  
Sequence: 1 gccagagattggcagcagg.....aattttgtattgtgtgtttt 6308.

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2359870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6308	100.0	6308	12	Adj95441 Human cDN
2	6308	100.0	6308	13	Ad86863 cDNA enco
3	5592	88.6	7742	8	ABZ24689 Human cel
4	5388	85.4	7648	12	Adi16377 Human pro
5	5136	81.4	5205	12	Adj95458 Human cDN
6	5136	81.4	5205	13	Ad86880 cDNA enco
7	4435.4	70.3	6395	2	Aav99308 cDNA enco
8	4435.4	70.3	6395	5	Aac86933 Nucleotid
9	3056.8	48.4	3059	10	Ada53348 Human cod
10	2546.8	40.4	2550	4	Aah14878 Human cDN
11	1716.4	27.2	3776	13	Acn39798 Tumour-as
12	1529.2	24.2	3598	13	Adr07317 Full leng
13	1381.8	21.9	6300	12	Adj95443 Human cDN
14	1381.8	21.9	6300	13	Ad86865 cDNA enco
15	1311	20.8	1635	6	Abq75898 Human ubi
16	1233.6	19.6	6089	12	Adj95445 Mouse cDN
17	1233.6	19.6	6089	13	Ad86867 cDNA enco
18	1099.4	17.4	6840	4	Aak51709 Human pol
19	1090	17.3	6850	4	Aak52693 Human pol
20	996.2	15.8	1001	2	Aax03300 Partial c

21	996.2	15.8	1001	5	Aac86934	Aac86934 Nucleotid
22	981	15.6	6158	13	ADR14633	ADR14633 Human NF-
23	813.6	12.9	3327	2	AAX35730	AAX35730 cDNA enco
24	813.6	12.9	3502	2	AAX35731	AAX35731 cDNA enco
25	797.2	12.6	4208	12	ADQ64363	Adq64363 Novel hum
26	656.8	10.4	733	3	AAA02411	Aaa02411 Human col
27	647	10.3	712	4	Aah07621	Aah07621 Human cDN
28	638.6	10.1	756	3	AAA02327	Aaa02327 Human col
29	626.8	9.9	2616	6	ABZ70161	Abz70161 Fibrinoge
30	450.8	7.1	777	5	AAS70068	Aas70068 DNA enco
31	438.4	6.9	505	9	ACH49509	Ach49509 Human leu
32	429	6.8	1464	5	AAS70104	Aas70104 DNA enco
33	374.8	5.9	1772	6	ABL90577	Abi90577 Human pol
34	294.4	4.7	4073	13	ADR07162	Adr07162 Full leng
35	271.4	4.3	1172	5	AAS70851	Aas70851 DNA enco
36	250.6	4.0	907	6	ABK35394	Abk35394 Human cDN
37	246.8	3.9	802	2	AAX39891	Aax39891 Gastric c
38	241.6	3.8	264	3	AAA00233	Aaa00233 Human col
39	219.8	3.5	4573	8	ACA03886	Acc03886 cDNA upre
40	216.8	3.4	6278	4	ABL14799	Abi14799 Drosophil
c 41	191.6	3.0	800	2	AAX39892	Aax39892 Gastric c
c 42	160.2	2.5	3411	10	ADB62029	Adb62029 Human cDN
c 43	159	2.5	455	2	AAV89674	Aav89674 EST clone
c 44	142	2.3	643	6	ABL89546	Abi89546 Human pol
c 45	136.6	2.2	550	12	ACH79702	Ach79702 Human gen

## ALIGNMENTS

RESULT 1  
ADJ95441

ID ADJ95441 standard; cDNA; 6308 BP.

AC ADJ95441;

DT 03-JUN-2004 (first entry)

DE Human cDNA encoding Ubiquitin ligase E3alpha I.

KW Human; ss; gene; ubiquitin ligase; E3alpha I;  
KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;  
KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;  
KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;  
KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;  
KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome;  
KW SNP; single nucleotide polymorphism.

OS Homo sapiens.

FH Key Location/Qualifiers  
CDS 696..5945

FT /\*tag= a

FT /\*product= "Ubiquitin ligase E3alpha I"

FT /\*replace(5397,C)

FT /\*tag= b

FT /\*standard\_name= "Single nucleotide polymorphism"

US6706505-B1.

PD 16-MAR-2004.

PF 28-NOV-2000; 2000US-00724126.

PR 08-MAR-2000; 2000US-0187911P.

PA (AMGE-) AMGEN INC.

PI Han H, Kwak K;

DR WPI; 2004-236723/22.

DR P-PsDB; ADJ95442.

XX



PT	New nucleic acid molecule, useful for preparing a composition for diagnosing, treating or preventing diseases associated with human E3approximateyla polypeptide, e.g., muscle atrophy.
PT	Claim 1; SEQ ID NO 1; 104pp; English.
XX	
PS	
CC	The invention relates to a new isolated nucleic acid molecule appearing as ADJ95441(or its complement) encoding a ubiquitin ligase E3alpha I protein appearing as ADJ95442. Also included are a vector comprising the nucleic acid, a host cell comprising the vector, a process of producing the E3alpha I ubiquitin ligase polypeptide, a composition comprising the nucleic acid molecule, a reagent comprising a detectably labelled nucleotide, and a method for determining the presence of a human E3alpha I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid molecule is useful for preparing a composition for diagnosing, treating or preventing diseases associated with human E3alpha I polypeptide, e.g. muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal cachexia, inflammatory cachexia, muscle wasting disorders associated with metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome, fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting syndrome. The present sequence encodes human E3alpha I.
CC	Sequence 6308 BP; 1856 A; 1301 C; 1390 G; 1761 T; 0 U; 0 Other;
XX	
Qy	Query Match 100.0%; Score 6308; DB 12; Length 6308;
Db	Best Local Similarity 100.0%; Pred. No. 0;
Qy	Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Db	1 GCCAAGAAATTCGGACGAGGGGAAAAGCTGAGCCGAGGACCAAAATTACTTCTTACCT 60
Qy	1 GCCAAGAAATTCGGACGAGGGGAAAAGCTGAGCCGAGGACCAAAATTACTTCTTACCT 60
Db	61 CATTTGTGAAGACAAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAACTTACCT 120
Qy	61 CATTTGTGAAGACAAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAACTTACCT 120
Db	121 CCAAAGGCTTATCATCTGTCTTCCACTTATCCAAACAAGCTGCTATGGCCACTGCCTGTGC 180
Qy	121 CCAAAGGCTTATCATCTGTCTTCCACTTATCCAAACAAGCTGCTATGGCCACTGCCTGTGC 180
Db	181 CGCAGCTGGAACACCGCCAGGCCCACTATCGCCTCCAGTACACTGGTTCGCCACCT 240
Qy	181 CGCAGCTGGAACACCGCCAGGCCCACTATCGCCTCCAGTACACTGGTTCGCCACCT 240
Db	241 GATCAGCTGCTGTGCTGCGCCATCTTATCCGCTTCTGCTGTTCTGAGTAAATGATACA 300
Qy	241 GATCAGCTGCTGTGCTGCGCCATCTTATCCGCTTCTGCTGTTCTGAGTAAATGATACA 300
Db	301 CACCTCGAAACACCACTTCTACTTTCTGTGCTATGAATTTGACTACTCTAGCTGGATC 360
Qy	301 CACCTCGAAACACCACTTCTACTTTCTGTGCTATGAATTTGACTACTCTAGCTGGATC 360
Db	361 CCGAGCTTTTTTGTACATGTCGAAGTGCCTCCACGGGGTAGAATCCTAAAAATAGAAAT 420
Qy	361 CCGAGCTTTTTTGTACATGTCGAAGTGCCTCCACGGGGTAGAATCCTAAAAATAGAAAT 420
Db	421 GTATGCAACAGTTCCAGCACCAACCCAGATATACAACATTCAGTACCAAGAGCTAC 480
Qy	421 GTATGCAACAGTTCCAGCACCAACCCAGATATACAACATTCAGTACCAAGAGCTAC 480
Db	481 GCCTGATAAAATTAGAGGGGAAAAAAAATCTCCAGTCCCTTTCAGTCCGTGACGCTTGCT 540
Qy	481 GCCTGATAAAATTAGAGGGGAAAAAAAATCTCCAGTCCCTTTCAGTCCGTGACGCTTGCT 540
Db	541 TCCGGGAAGCGGGCCGGAAGCCACTCTCGAGTCTCGGTCAAAACCCGACTTCAGGGGCGG 600
Qy	541 TCCGGGAAGCGGGCCGGAAGCCACTCTCGAGTCTCGGTCAAAACCCGACTTCAGGGGCGG 600
Db	601 TCGTAAAGTGTGTCCTGCTCTCCGACCGGGCCACAGGTTTCGCTTCGCTCTGCGCG 660
Qy	601 TCGTAAAGTGTGTCCTGCTCTCCGACCGGGCCACAGGTTTCGCTTCGCTCTGCGCG 660
Db	661 GGGGTCCGCAACTGACGGCGTCAGTTTCCCTCAAGATGGCGGACGAGGAGGCTGGAGGTA 720
Qy	661 GGGGTCCGCAACTGACGGCGTCAGTTTCCCTCAAGATGGCGGACGAGGAGGCTGGAGGTA 720



QY	3961	CTGTTACTGAAAGGAGGTGCTGACGTGCATCTTTGCAAGAAACAGAGGTTGAAA	4020
Db	3961		
QY	4021	TAGAAAATAATGCCATGCTGATTTATCGGCTGTGTCAGAAATCTACTGCCTTAACCCAGC	4080
Db	4021		
QY	4081	ACAGGGGAAAACCCATAGAACTCTCAGAGAGCCCTAGACCCTTTTCATGATCCAG	4140
Db	4081		
QY	4141	ACTTGGCATATGGAATCTTATACAGAAAGCTGTGTCATGTAATGCAACGAGTGTCTGC	4200
Db	4141		
QY	4201	AGAACTATTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCANGTGACCTTTTGTG	4260
Db	4201		
QY	4261	ACTTGGAAAGTGAGAAATATCTTTGCCCTCTTTGCAAAATCTCTGTGCAATCTGTGATCC	4320
Db	4261		
QY	4321	CCATTTATCTCTTTGCAACTCAAAAGATAAACAGTGAGATGCGATGCTCTTGCTCAAC	4380
Db	4321		
QY	4381	TTTTGACCTCTGACGCTGGATACAGACTGTTCTGGCCAGAAATATCAGGTTTATATATAA	4440
Db	4381		
QY	4441	GACATGCTAAAGGAGAAACCCAAATTCCTATTTCTTTAAATCAAGGAATGGAGATTCTA	4500
Db	4441		
QY	4501	CTTTGGAGTTCATTTCCATCTGAGTCTTTGGGTTGAGTCTTCGATTAATAATTCAAATA	4560
Db	4501		
QY	4561	GCATCAAGGAATGGTTATCTCTTTGCCACAAATTTATAGAAATGGATTGAAAGTGC	4620
Db	4561		
QY	4621	CACCTGATGAAAGGATCTCGAGTCCCAATGCTGAGTCCCAATGCTGAGAGCACCTGCGCTTTCACTA	4680
Db	4621		
QY	4681	TCCAGGCAATTCAAATCTATTGGGAGATGAAGGAACCTCTGTTTGGAGCACCTTCAAA	4740
Db	4681		
QY	4741	ATAGGCAGCAATAATGGTCTGAAAGCAATTAATGCAGTTTGCAGTTGTCACAGAGGATTACCT	4800
Db	4741		
QY	4801	GTCTCAGTCTCTGATACAGAAACATCTGGTTGCTCTCTATACAGTTGTTCTTCTCAACA	4860
Db	4801		
QY	4861	TAAATCAGAGATACACCATGCTCTCTATAGATCTGTTTCAATGTTTGGTGGGTG	4920
Db	4861		
QY	4921	CTGTGTAGCATTTCCCATCTTTGATTTGGGATGACCCCTGTTGATCTGCAGCCCTTCTCAG	4980
Db	4921		
QY	4981	TTAGTTCTTCTTATACCACTTTTATCTCTTCCATTTGATACCAATGGCACAATGCTTC	5040
Db	4981		
QY	5041	AGATACTACTTACAGTACAGACAGGCTTACCCCTTGCTCAGGTTCAAGAGACAGTGAAG	5100
Db	5041		

Db	5041	AGATACTACTTACAGTACAGACAGGCTTACCCCTTGCTCAGGTTCAAGAGACAGTGAAG	5100
QY	5101	AGGCTCATTCGCGATCTTCTTTCTTTGCGAGAAATTTCTCAATATACAGTGGCTCCATTG	5160
Db	5101		
QY	5161	GGTGTGATATTCTCGGCTGGTATTGTTGGGTCTCACATGAAGAAATGGCATCACCCCTTATC	5220
Db	5161		
QY	5221	TTCTCTGTCTGTCATTTGTTTTCACATATTACTTTGGGTAACTCCGCTGAGGAACTGC	5280
Db	5221		
QY	5281	ATACCAATTTCTGCAAGAGGAGTACAGTGCACTCTGTAGCTATCTATCTTACTACAA	5340
Db	5281		
QY	5341	ATTTGTTCTGCTCTCCAGGAATATTGGGATACGTAAAGCCCTTGCTCCAGAGGTGGT	5400
Db	5341		
QY	5401	GTGAGATCTCTGCTTACTTAACTGTTTGAAGCAAAAAACAACCGTGGTCAAGTACCCTA	5460
Db	5401		
QY	5461	GAAAAAGAAATAGTTTGTATAGAGCTTCTGTAGCTATAGCTGCTCTGAATCAAGCTT	5520
Db	5461		
QY	5521	CTCATTTTCAGTGCCACGGTCTGCAGATGATGAGCAAGCATCTGCTCTGCTGCTTTT	5580
Db	5521		
QY	5581	TCCTGCGGCTATATCTATGTTCTCAGAACTTTCTGCGCAGGAAATTTGTAAACCGGGAG	5640
Db	5581		
QY	5641	AGTTTGGAGCTTGCAATTTTTCACGCACTTCACTGTGAGCGCGAGTCTGCAATTTTCTTAA	5700
Db	5641		
QY	5701	AAATCAGAGATGCCGAGTGTCTGTTGAAGTAAAGCCAGAGGCTGTGCTCATCCAG	5760
Db	5701		
QY	5761	CTCCTTACTTGGATGAATATGGAGAAACAGACCCCTGCGCTGAAGAGGGCAACCCCTTTC	5820
Db	5761		
QY	5821	ATTTATCTGTGAGGGTATCGGAAGCTCCATTTGGTCTGGCAACAACATCTGCATTATAG	5880
Db	5821		
QY	5881	AAGGATTTCTAGAGCCCAAGAGACTAATCAGATGTTATTTGGATTCAACTGGCAGTTAC	5940
Db	5881		
QY	5941	TGTGAGCTTCAACTCTGCTCAAGACAATCAAAATGACGACAGTAGTAAAGGCTGATTC	6000
Db	5941		
QY	6001	AAAAATATGGAATCTTTCTGAGGCTGGGAAAGTATTTGGAGGTCTTTTCTCTCATGTC	6060
Db	6001		
QY	6061	CAGGTTCACTTACATCAATAAATATTTCTTAAATGGAGTATTTGCTTCAATAGCAACA	6120
Db	6061		
QY	6121	TATGCTTACAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA	6180
Db	6121		

Db 6121 TATGCTTACAGGAAAGGACATAGATCAATCTGTTTATGTGCTAGTATTTCCAGGA 6180  
 QY 6181 ATTATTCCTTCAATAATTGCTCATTTTATTTATTCATCCACTTGGTAGATGAAG 6240  
 Db 6181 ATTATTCCTTCAATAATTGCTCATTTTATTTATTCATCCACTTGGTAGATGAAG 6240  
 QY 6241 TCACGTCAACAGTGTAGACATTTTATGCTGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300  
 Db 6241 TCACGTCAACAGTGTAGACATTTTATGCTGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300  
 QY 6301 GGTGTTTT 6308  
 Db 6301 GGTGTTTT 6308

## RESULT 2

ADS86863

ID ADS86863 standard; cDNA; 6308 BP.

XX AC ADS86863;

XX DT 16-DEC-2004 (first entry)

XX DE cDNA encoding human E3alpha ubiquitin ligase, hE3alphaI.

XX KW Human; E3alpha ubiquitin ligase; huE3alphaI;

XX KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;  
 KW fasting; metabolic acidosis; muscle degeneration; kidney failure;  
 KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;  
 KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;  
 KW inflammatory cachexia; hyperthyroidism; denervation atrophy;  
 KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;  
 KW gene therapy; ss; gene.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT 5'UTR 1..695

FT /\*tag= a

FT CDS 696..5945

FT /\*tag= b

FT /product= "E3alphaI"

FT 3'UTR 5946..6308

FT /\*tag= c

XX PN US2004185037-A1.

XX PD 23-SEP-2004.

XX PF 15-JAN-2004; 2004US-00758672.

XX PR 08-MAR-2000; 2000US-0187911P.

XX PR 28-NOV-2000; 2000US-00724126.

XX XX (HANH/) HAN H.

XX PA (KWAK/) KWAK K.

XX PI Han H, Kwak K;

XX DR WPI; 2004-707854/69.

XX DR P-P5DB; ADS86864.

XX PT Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule  
 XX PT useful for treating and/or preventing renal cachexia or inflammatory  
 XX PT cachexia.

XX PS Claim 1; SEQ ID NO 1; 115pp; English.

XX CC The present invention relates to new orthologue of human E3alpha  
 XX CC ubiquitin ligase, huE3alphaI and huE3alphaII. Most intracellular proteins  
 XX CC are degraded through the ubiquitin-proteasome pathway. Proteins are  
 XX CC marked for proteasomal degradation by conjugation of ubiquitin to the  
 XX CC protein. Conjugation of the ubiquitin molecule involves the activation by

CC E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,  
 CC and then interacts with a specific E3 ligase family member. E3 ligase  
 CC binds to proteins targeted for degradation and catalyses the transfer of  
 CC ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase  
 CC determines the specificity of the system. The E3alpha family is the main  
 CC family of intracellular ligases and is involved in the N-end rule pathway  
 CC of protein degradation. E3alpha enzyme binds directly to the primary  
 CC destabilising N-terminal amino acid and catalyses ubiquitin conjugation  
 CC thereby targeting the protein for degradation. The human E3alpha gene is  
 CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-  
 CC proteasome pathway has been determined to be a major cause of rapid  
 CC muscle wasting including, fasting, metabolic acidosis, muscle  
 CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,  
 CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen  
 CC balance, burns, Cushing's syndrome, inflammatory cachexia,  
 CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-  
 CC protein malnutrition. E3alpha plays a role in the overall increase in  
 CC ubiquitination that is associated with and may mediate muscle atrophy in  
 CC catabolic and other disease states. Treatment may be administered by gene  
 CC therapy, cell therapy and antisense therapy methods. The present sequence  
 CC is cDNA encoding human E3alphaI.

XX SQ Sequence 6308 BP; 1856 A; 1301 C; 1390 G; 1761 T; 0 U; 0 Other;

Query Match 100.0%; Score 6308; DB 13; Length 6308;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAAGAATTCGGCAGGAGGGGAAAGCTGAGCCAGGACCAATTAATTTACTTTCCT 60

Db 1 GCCAAGAATTCGGCAGGAGGGGAAAGCTGAGCCAGGACCAATTAATTTACTTTCCT 60

QY 61 CATTTGTGAAGACAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAGAACTTACCT 120

Db 61 CATTTGTGAAGACAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAGAACTTACCT 120

QY 121 CCAAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGTATGGCCACTGCTGTGC 180

Db 121 CCAAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGTATGGCCACTGCTGTGC 180

QY 181 CGCAGCTGGAAACACCGCCAGCCCACTACTGCGCTCCCACTGCTGTTCTCCACCCCT 240

Db 181 CGCAGCTGGAAACACCGCCAGCCCACTACTGCGCTCCCACTGCTGTTCTCCACCCCT 240

QY 241 GATCAGCTGCTGTGCTGCCATCTTATCCGCTTCTGCTGCTTCTGAGTAAATGTATACA 300

Db 241 GATCAGCTGCTGTGCTGCCATCTTATCCGCTTCTGCTGCTTCTGAGTAAATGTATACA 300

QY 301 CACCTTGAAACACCGCCAGCCCACTACTGCGCTCCCACTGCTGTTCTCCACCCCT 360

Db 301 CACCTTGAAACACCGCCAGCCCACTACTGCGCTCCCACTGCTGTTCTCCACCCCT 360

QY 361 CCGAGCTTTTGTACACATGTGCAAGTCCCAAGGCTAGATCTCTAAATAGAGAT 420

Db 361 CCGAGCTTTTGTACACATGTGCAAGTCCCAAGGCTAGATCTCTAAATAGAGAT 420

QY 421 GTATGCAACAGTTCAGCAGCAAAACCCAGATATACACCAATTCAGCTACCAAGAGCTAC 480

Db 421 GTATGCAACAGTTCAGCAGCAAAACCCAGATATACACCAATTCAGCTACCAAGAGCTAC 480

QY 481 GCCTGATAAATTAGAGGGGAAAAAATCTCCAGTCCCTTTCAGCTCGTGAAGCTTGC 540

Db 481 GCCTGATAAATTAGAGGGGAAAAAATCTCCAGTCCCTTTCAGCTCGTGAAGCTTGC 540

QY 541 TCCGGAGAGCGGCGGAGGACCTCTCGAGTCTGCGTCAAAACCCGACTTCAGGGGCG 600

Db 541 TCCGGAGAGCGGCGGAGGACCTCTCGAGTCTGCGTCAAAACCCGACTTCAGGGGCG 600

QY 601 TCGTAAAGTCTGCTCCCTGCTCTCCGACCGGCGCAGAGTTTCCGCTTTCCTCTGCGCG 660

Db 601 TCGTAAAGTCTGCTCCCTGCTCTCCGACCGGCGCAGAGTTTCCGCTTTCCTCTGCGCG 660

QY 661 GGGGTGCGCAACTGCGAGCGCTCAGTTTCCCTCAAGATGCGCGGAGGAGGCTGAGGTA 720

Db	661	GGGGTGGGCACTGAGGGTTCAGTTCCCTCAAGATGGGGAGGGGCTGGAGGT	720
Qy	721	CTGAGAGATGGAATTCAGCGGGAGTTACCCAGAGCCCTCAGGGTCTGGCATCTTGGT	780
Db	721	CTGAGAGATGGAATTCAGCGGGAGTTACCCAGAGCCCTCAGGGTCTGGCATCTTGGT	780
Qy	781	GGGATCAGCAGTTGATTTTATACCTTTCTTCATCATTTGGCAAAATGGTGCCAG	840
Db	781	GGGATCAGCAGTTGATTTTATACCTTTCTTCATCATTTGGCAAAATGGTGCCAG	840
Qy	841	AAATTTACTTTGCTGAAATGACCCAGACTTGGAAAAGCAGGAGAAAGTGTACAAATGT	900
Db	841	AAATTTACTTTGCTGAAATGACCCAGACTTGGAAAAGCAGGAGAAAGTGTACAAATGT	900
Qy	901	CAATATTCATCCTCACTGGAATGGTACTTATTTGGAGAAAGATCCAGATATTTGCTTAGAGA	960
Db	901	CAATATTCATCCTCACTGGAATGGTACTTATTTGGAGAAAGATCCAGATATTTGCTTAGAGA	960
Qy	961	AAATGAAGCAGAGTGGAGCAATTCAGCTTTGGGAGGGTTCAAAAGTGGAGAGACAA	1020
Db	961	AAATGAAGCAGAGTGGAGCAATTCAGCTTTGGGAGGGTTCAAAAGTGGAGAGACAA	1020
Qy	1021	CCTATTTCTGAGGATTTGCAATTCATTCGAAATGTGTACTCTGTATGGAATCTCTCC	1080
Db	1021	CCTATTTCTGAGGATTTGCAATTCATTCGAAATGTGTACTCTGTATGGAATCTCTCC	1080
Qy	1081	AGGACAGTGTTCATAAAATCATCGTTACAAGATGCATATTTCTACTGGAGGAGGTCT	1140
Db	1081	AGGACAGTGTTCATAAAATCATCGTTACAAGATGCATATTTCTACTGGAGGAGGTCT	1140
Qy	1141	GTGACTGTGGAGACACAGAGGATTCGAAATTCGAAATGTGTACTCTGTATGGAATCTCTCC	1200
Db	1141	GTGACTGTGGAGACACAGAGGATTCGAAATTCGAAATGTGTACTCTGTATGGAATCTCTCC	1200
Qy	1201	GAAGAGCAGTACTATAAAGAGAAATTCAGCGTTCGTTGAATCAAGAGGTAAATGTCTCC	1260
Db	1201	GAAGAGCAGTACTATAAAGAGAAATTCAGCGTTCGTTGAATCAAGAGGTAAATGTCTCC	1260
Qy	1261	AAAGCAGGAAAATATTTCTTCAGTGTATAAATAATGTGTAGAAATGCTATATGGGAAG	1320
Db	1261	AAAGCAGGAAAATATTTCTTCAGTGTATAAATAATGTGTAGAAATGCTATATGGGAAG	1320
Qy	1321	AGGAAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAATCAAGATATTAATTTGTG	1380
Db	1321	AGGAAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAATCAAGATATTAATTTGTG	1380
Qy	1381	TCCTTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGCCCTACAAAGAGCTC	1440
Db	1381	TCCTTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGCCCTACAAAGAGCTC	1440
Qy	1441	TTGACTGTGAGTCGCGAGGCCCAGTTGTCATACCACTGCTTATACAGAGGGTCTGTC	1500
Db	1441	TTGACTGTGAGTCGCGAGGCCCAGTTGTCATACCACTGCTTATACAGAGGGTCTGTC	1500
Qy	1501	GGGCTGTAAAGCGGGAGTTATGCTTTCGAGGACCAAGCAAGATATAAAGAGCTC	1560
Db	1501	GGGCTGTAAAGCGGGAGTTATGCTTTCGAGGACCAAGCAAGATATAAAGAGCTC	1560
Qy	1561	ATTGAGAAATGCTCTCAACATCCACTTCATGTAGAGTATTAACATCAGAGATTTATGG	1620
Db	1561	ATTGAGAAATGCTCTCAACATCCACTTCATGTAGAGTATTAACATCAGAGATTTATGG	1620
Qy	1621	CTCATCAGAAATTTGCTTGGCTCTTGGTCTCTGGATGAACAAAATTTATGACTATTCAA	1680
Db	1621	CTCATCAGAAATTTGCTTGGCTCTTGGTCTCTGGATGAACAAAATTTATGACTATTCAA	1680
Qy	1681	GTGACTTTAGGAGATCTTTTGGCAAGCATGCTTAGAGAGAACTTGACTCGGAGATC	1740
Db	1681	GTGACTTTAGGAGATCTTTTGGCAAGCATGCTTAGAGAGAACTTGACTCGGAGATC	1740
Qy	1741	CTGTCTCATAAAGCAGTTAAATGCTTTGGGATGCAAGCTTTTAAAGGTGCCCGCTAAGA	1800

Db	1741	CCTCTCTATAAGCAGGTTAAATGCTTTGGGATGCAAAAGCTTTATAAAAGGTGCCCGCTAAGA	1800
Qy	1801	TCCTTTCATGAATTCATCTTCAGCAGTTTTCATTTTATGAGAGATGGAATACAAAACCTCTTTG	1860
Db	1801	TCCTTTCATGAATTCATCTTCAGCAGTTTTCATTTTATGAGAGATGGAATACAAAACCTCTTTG	1860
Qy	1861	CTATGGAATTTGTGAAGTTATATAAACAACCTGCAAGAAAGATATATCAGTGATGATCATG	1920
Db	1861	CTATGGAATTTGTGAAGTTATATAAACAACCTGCAAGAAAGATATATCAGTGATGATCATG	1920
Qy	1921	ACAGAAGTATCTCTATATACTGCACTTCCTCAGTTCAGATGTTTATCTGTTCTTACTCTGGCTC	1980
Db	1921	ACAGAAGTATCTCTATATACTGCACTTCCTCAGTTCAGATGTTTATCTGTTCTTACTCTGGCTC	1980
Qy	1981	GACATCTTATGAGAGCAGAACTGTTATCTCTGTCATTTACTGAAACTCTGCTAGAGTTT	2040
Db	1981	GACATCTTATGGAAGAGCAGAACTGTTATCTCTGTCATTTACTGAAACTCTGCTAGAGTTT	2040
Qy	2041	TACCTGAGTACTTCGGAAGGAAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAAT	2100
Db	2041	TACCTGAGTACTTCGGAAGGAAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAAT	2100
Qy	2101	TGGGAAGATATATGCAAGTAAATATGTCGACCTTAAAGTATATCTCTGATCAGCAAAACCAAA	2160
Db	2101	TGGGAAGATATATGCAAGTAAATATGTCGACCTTAAAGTATATCTCTGATCAGCAAAACCAAA	2160
Qy	2161	TATGGAAGAAAGATTAAGATGCACTTCCTTGAAGTTTTCGATCTTTTGAAGATTC	2220
Db	2161	TATGGAAGAAAGATTAAGATGCACTTCCTTGAAGTTTTCGATCTTTTGAAGATTC	2220
Qy	2221	TTACTGCTGATGCAAGGAAATGGAAGAAATCCGAAGCAGTTGGGCAACACATTTGAAGTGG	2280
Db	2221	TTACTGCTGATGCAAGGAAATGGAAGAAATCCGAAGCAGTTGGGCAACACATTTGAAGTGG	2280
Qy	2281	ATCTGATTTGGAGGCTGCCATTCCTATACAGATGCAATTCGAAGAAATTTTACTCATGT	2340
Db	2281	ATCTGATTTGGAGGCTGCCATTCCTATACAGATGCAATTCGAAGAAATTTTACTCATGT	2340
Qy	2341	TCCAAGATGCTGCTTGTGATGCAAGAACTTCCTTACTTGGCTTTATAAAGATGTCACA	2400
Db	2341	TCCAAGATGCTGCTTGTGATGCAAGAACTTCCTTACTTGGCTTTATAAAGATGTCACA	2400
Qy	2401	AAGCTGCTGATGAGGTGCACTGATTCATATCTAGTAGCAAGACAGTAGTACAAATCGT	2460
Db	2401	AAGCTGCTGATGAGGTGCACTGATTCATATCTAGTAGCAAGACAGTAGTACAAATCGT	2460
Qy	2461	GTGGAATGATTTTGGAAAACAAAGTCCACAGATATCTGAGGATCTTGTAGGACATATC	2520
Db	2461	GTGGAATGATTTTGGAAAACAAAGTCCACAGATATCTGAGGATCTTGTAGGACATATC	2520
Qy	2521	TGCCACTCTCTAGGACCTTCTGCTCTCATGTACGTTTAAAGCAGGCTGGTCTGTTT	2580
Db	2521	TGCCACTCTCTAGGACCTTCTGCTCTCATGTACGTTTAAAGCAGGCTGGTCTGTTT	2580
Qy	2581	CAAGACTGCAATTTTGTGCTTTTGAAGGACTTTTCAAGTAGAGGTACTAGTGGAAATATC	2640
Db	2581	CAAGACTGCAATTTTGTGCTTTTGAAGGACTTTTCAAGTAGAGGTACTAGTGGAAATATC	2640
Qy	2641	CTTTACGTTGCTGGTGTGGTCCAGGTTGCTGAGATGTCGAGATGTCGGAAGAAATGGAC	2700
Db	2641	CTTTACGTTGCTGGTGTGGTCCAGGTTGCTGAGATGTCGAGATGTCGGAAGAAATGGAC	2700
Qy	2701	TGCTCTTTATGAGCAGGTTTATTAACCAAGATGTTAAGTGCAGAGAAATGATATG	2760
Db	2701	TGCTCTTTATGAGCAGGTTTATTAACCAAGATGTTAAGTGCAGAGAAATGATATG	2760
Qy	2761	ATAAAGATATCATCATGCTTCAGATTCGTGATCTTTTAAATGGATCCCAATAGTTCTTGT	2820
Db	2761	ATAAAGATATCATCATGCTTCAGATTCGTGATCTTTTAAATGGATCCCAATAGTTCTTGT	2820
Qy	2821	TACTGGTATCTTCAGAGGTATGAACCTTGGCGAGGCTTTTAAACAGACCATATCTACAAAG	2880
Db	2821	TACTGGTATCTTCAGAGGTATGAACCTTGGCGAGGCTTTTAAACAGACCATATCTACAAAG	2880



QY 2881 ACCAGGATTTGATTAACAAATATATATACATTAATAGAAAGAAATGCTTCAGGTCCTCATCT 2940  
DB 2881 ACCAGGATTTGATTAACAAATATATATACATTAATAGAAAGAAATGCTTCAGGTCCTCATCT 2940  
QY 2941 ATATTGTCGGGTGAGCGTTATGTACCTGAGGTGGGAAATGTGACCAAAAGAGAGTCAACA 3000  
DB 2941 ATATTGTCGGGTGAGCGTTATGTACCTGAGGTGGGAAATGTGACCAAAAGAGAGTCAACA 3000  
QY 3001 TGAGAGAAATCAATTCATCTGCTTTGCAATTGAACCCATGCCACACAGTGCCTTGCACAAA 3060  
DB 3001 TGAGAGAAATCAATTCATCTGCTTTGCAATTGAACCCATGCCACACAGTGCCTTGCACAAA 3060  
QY 3061 ATTTACTGAGAAATGAATAATTAATGAACCTGGCTTAGAGAAATGTCTATAAAAGAGTGGCCA 3120  
DB 3061 ATTTACTGAGAAATGAATAATTAATGAACCTGGCTTAGAGAAATGTCTATAAAAGAGTGGCCA 3120  
QY 3121 CATTTAAGAAACCAAGTGTATCAGGCCATGGAGTTTATGAACCTAAGAGTGAATCACTGA 3180  
DB 3121 CATTTAAGAAACCAAGTGTATCAGGCCATGGAGTTTATGAACCTAAGAGTGAATCACTGA 3180  
QY 3181 AAGACTTCAATATGTACTTTTATCAATTAATCAAAACCCAGCATAGCAAGGCTGAACATA 3240  
DB 3181 AAGACTTCAATATGTACTTTTATCAATTAATCAAAACCCAGCATAGCAAGGCTGAACATA 3240  
QY 3241 TGCAGAGAAAGAGGAGAAACCAAGAAACAAAGATGAAGCAATGCGGCCACACCACTC 3300  
DB 3241 TGCAGAGAAAGAGGAGAAACCAAGAAACAAAGATGAAGCAATGCGGCCACACCACTC 3300  
QY 3301 CTGAATTCGCGCTCTTCAGCAAGTGAATTAACCTTCTCACTGTGATATCATGATGT 3360  
DB 3301 CTGAATTCGCGCTCTTCAGCAAGTGAATTAACCTTCTCACTGTGATATCATGATGT 3360  
QY 3361 ACATTTCTCAGAACCGTATTTGAGCGGGCAATAGACACAGATTTCAACTTGTGGACCGAAG 3420  
DB 3361 ACATTTCTCAGAACCGTATTTGAGCGGGCAATAGACACAGATTTCAACTTGTGGACCGAAG 3420  
QY 3421 GGATGCTCCAAATGGCTTTTCAATTTCTGGCAATGGGTTTACTAGAGAGAAAGCAACAGC 3480  
DB 3421 GGATGCTCCAAATGGCTTTTCAATTTCTGGCAATGGGTTTACTAGAGAGAAAGCAACAGC 3480  
QY 3481 TTCAAAAAGCTCTCCTGAAGAGAAATCAATTTGACTTTTATCATNAAGGCTTCAAGATTGG 3540  
DB 3481 TTCAAAAAGCTCTCCTGAAGAGAAATCAATTTGACTTTTATCATNAAGGCTTCAAGATTGG 3540  
QY 3541 GAAGTTTCAAGCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTCGCCAGTTAG 3600  
DB 3541 GAAGTTTCAAGCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTCGCCAGTTAG 3600  
QY 3601 AAGGCCAGAGGACATGATTAACGTGGATACCTTCAGATGCTTTGACACAGTGAAGCGATTAA 3660  
DB 3601 AAGGCCAGAGGACATGATTAACGTGGATACCTTCAGATGCTTTGACACAGTGAAGCGATTAA 3660  
QY 3661 GAGAAAAATCTTGTTTAAATTTAGCAACCACTCAGGATCGGAATCTAATGAAGATGATG 3720  
DB 3661 GAGAAAAATCTTGTTTAAATTTAGCAACCACTCAGGATCGGAATCTAATGAAGATGATG 3720  
QY 3721 AGATTACTCATGATAAAGAAAGCAGAAACGAAAGAAAGAGCTGAAGCTCTAGGCTTAC 3780  
DB 3721 AGATTACTCATGATAAAGAAAGCAGAAACGAAAGAAAGAGCTGAAGCTCTAGGCTTAC 3780  
QY 3781 ATCGCAGAGAGATCATGCTCAGATGCTGCTGCTTACAGAAAACTTCAATTTGAACCTATA 3840  
DB 3781 ATCGCAGAGAGATCATGCTCAGATGCTGCTGCTTACAGAAAACTTCAATTTGAACCTATA 3840  
QY 3841 AACTCATGTATGAACAATCATCAGAAATGCTGGGAAAGAGATTTCCATTTATGAGGAG 3900  
DB 3841 AACTCATGTATGAACAATCATCAGAAATGCTGGGAAAGAGATTTCCATTTATGAGGAG 3900  
QY 3901 AGAGACCCAGCAGTCACTGCTACTAGAAATGCTTTGGTCTTAAACGGGGTCCAT 3960  
DB 3901 AGAGACCCAGCAGTCACTGCTACTAGAAATGCTTTGGTCTTAAACGGGGTCCAT 3960

QY 3961 CTGTTACTGAAAAAGAGGTGCTGACGTGCATCTCTTGGCCAAAGAAAGACAGAGGTGAAAA 4020  
DB 3961 CTGTTACTGAAAAAGAGGTGCTGACGTGCATCTCTTGGCCAAAGAAAGACAGAGGTGAAAA 4020  
QY 4021 TAGAAAAATATGCAATGGTATTTATCGGCTCTGTGTCAGAAATCTACTGCCCTTAACCCAGC 4080  
DB 4021 TAGAAAAATATGCAATGGTATTTATCGGCTCTGTGTCAGAAATCTACTGCCCTTAACCCAGC 4080  
QY 4081 ACAGGGGAAAAACCATAGAACTCTCAGGAGAGCCCTTAGACCCACTTTTTCATGGAATCCAG 4140  
DB 4081 ACAGGGGAAAAACCATAGAACTCTCAGGAGAGCCCTTAGACCCACTTTTTCATGGAATCCAG 4140  
QY 4141 ACTTTGGCATATGGAACCTTATACAGAGAGCTGTGCTCATGTAATGCACGAGTGTCTGGC 4200  
DB 4141 ACTTTGGCATATGGAACCTTATACAGAGAGCTGTGCTCATGTAATGCACGAGTGTCTGGC 4200  
QY 4201 AGAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260  
DB 4201 AGAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260  
QY 4261 ACTTTGGAAGTGGAGAAATATCTTTGCCCTCTTTTGCAAATCTCTGTGCAATATCTGTGATCC 4320  
DB 4261 ACTTTGGAAGTGGAGAAATATCTTTGCCCTCTTTTGCAAATCTCTGTGCAATATCTGTGATCC 4320  
QY 4321 CCATTTATCTTTTGCACCTCAAAAGATAAACAAGTGAAGTGCAGATGCTCTTGTCTCAAC 4380  
DB 4321 CCATTTATCTTTTGCACCTCAAAAGATAAACAAGTGAAGTGCAGATGCTCTTGTCTCAAC 4380  
QY 4381 TTTTGGCCCTGGCAGCGTGGATACAGACTGTTCTGGCCAGAAATATCAGGTTATATATAA 4440  
DB 4381 TTTTGGCCCTGGCAGCGTGGATACAGACTGTTCTGGCCAGAAATATCAGGTTATATATAA 4440  
QY 4441 GACATGCTAAAGAGAGAAACCCAAATTCCTATTTTCTTTTAAATCAAGGAATGGGAGATTCTA 4500  
DB 4441 GACATGCTAAAGAGAGAAACCCAAATTCCTATTTTCTTTTAAATCAAGGAATGGGAGATTCTA 4500  
QY 4501 CTTTGGAGTTCCTATTCATCTGAGTGTGCGGTGAGTCTTCGATTTAAATATTCAAATA 4560  
DB 4501 CTTTGGAGTTCCTATTCATCTGAGTGTGCGGTGAGTCTTCGATTTAAATATTCAAATA 4560  
QY 4561 GCATCAAGGAATGTTTATCTCTTGCACAAATTTATAGAAATGGATTTGAAGTGC 4620  
DB 4561 GCATCAAGGAATGTTTATCTCTTGCACAAATTTATAGAAATGGATTTGAAGTGC 4620  
QY 4621 CACCTGATGAAGGAGTCTCGAGTCCCATCTGACCTGGAGCACTTCGCGCTTCACTA 4680  
DB 4621 CACCTGATGAAGGAGTCTCGAGTCCCATCTGACCTGGAGCACTTCGCGCTTCACTA 4680  
QY 4681 TCCAGGCAATTTGAAAAATCTATTTGGGAGATGAAGAAAAACCTCTGTTTGGAGCACTTCAAA 4740  
DB 4681 TCCAGGCAATTTGAAAAATCTATTTGGGAGATGAAGAAAAACCTCTGTTTGGAGCACTTCAAA 4740  
QY 4741 ATAGGCGAGCAATATGCTGCAAGCATTTAATGCAAGTTTGCAGTTGTCACAGAGGATTACCT 4800  
DB 4741 ATAGGCGAGCAATATGCTGCAAGCATTTAATGCAAGTTTGCAGTTGTCACAGAGGATTACCT 4800  
QY 4801 GTCCTCAGGTCCTCATACAGAAAAATCTGCTTCTCTATCATGTTGTTCTTCTCTAACA 4860  
DB 4801 GTCCTCAGGTCCTCATACAGAAAAATCTGCTTCTCTATCATGTTGTTCTTCTCTAACA 4860  
QY 4861 TAAATCAGAGATACACCATGCTTCTGTCTATAGATCTGTTTCATGTTTGGTGGTG 4920  
DB 4861 TAAATCAGAGATACACCATGCTTCTGTCTATAGATCTGTTTCATGTTTGGTGGTG 4920  
QY 4921 CTGTGTTAGCAATTCCTGATTTGGGATGACCTGTTGATCTGCAGGCTTCTTCAG 4980  
DB 4921 CTGTGTTAGCAATTCCTGATTTGGGATGACCTGTTGATCTGCAGGCTTCTTCAG 4980  
QY 4981 TTAGTCTTCTCTATAACCACTTTTATCTTCCATTTTGTATCCATGTCACATGCTTC 5040  
DB 4981 TTAGTCTTCTCTATAACCACTTTTATCTTCCATTTTGTATCCATGTCACATGCTTC 5040  
QY 5041 AGATACTACTTACAGTATAGACAGGCTTACCCCTTGTCTCAGGTTTCAAGAGACAGTGAAG 5100



PT death, useful for treating, diagnosing or preventing cancer,  
PT developmental, neurological, reproductive or autoimmune/inflammatory  
XX disorders.

PS Claim 5; Page 216-218; 238pp; English.

XX The present sequence is that of Incyte clone 1351608C81 encoding human  
CC CGD-1, a novel protein associated with cell growth, differentiation and  
CC death. A representative cDNA library for the polynucleotide is P6ANN01  
CC from paragonomic tumour tissue. Structural features establish the  
CC encoded protein as being associated with cell growth, differentiation and  
CC death, with further evidence suggesting it to be a ubiquitin protein  
CC ligase. The invention is based on novel human CGD-1 to -21 proteins (see  
CC ABP5830-50), the polynucleotides encoding them (see ABZ4689-709), and  
CC to the use of these for the diagnosis, treatment or prevention of cell  
CC proliferative disorders including cancer, developmental disorders,  
CC neurological disorders, autoimmune disorders, reproductive disorders, and  
CC disorders of the placenta, and in the assessment of the effects of  
CC exogenous compounds on the activity and expression of proteins and  
CC nucleic acids associated with cell growth, differentiation and death.  
CC CGD polynucleotides are also used in a claimed microarray and in a  
CC claimed method of generating an expression profile of a sample

XX SQ Sequence 7742 BP; 2388 A; 1423 C; 1622 G; 2309 T; 0 U; 0 Other;

Query Match 88.6%; Score 5592; DB 8; Length 7742;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 717 GGTACTGAGGATGGAATCAGCGGAGTTACCCAGACCCCTCAGGCTGCGATCT 776  
DB 1 GGTACTGAGGATGGAATCAGCGGAGTTACCCAGACCCCTCAGGCTGCGATCT 60

QY 777 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGCAATCTTTGGCAATTTGGTG 836  
DB 61 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGCAATCTTTGGCAATTTGGTG 120

QY 837 CCAGAAATTTACTTTGCTGAAATGACCCAGACTTTGGAAAGCAGGAGAAAGTGTACAA 896  
DB 121 CCAGAAATTTACTTTGCTGAAATGACCCAGACTTTGGAAAGCAGGAGAAAGTGTACAA 180

QY 897 ATGTCAATTTCACTCCACTGGAAATGTTACTTTTGGAGAGATCCAGATATTTGCTTA 956  
DB 181 ATGTCAATTTCACTCCACTGGAAATGTTACTTTTGGAGAGATCCAGATATTTGCTTA 240

QY 957 GAGAAATTTGAGCAGTGGAGCATTTTCTGCTTTGCGAGGTTTCAAAGTGGAGAG 1016  
DB 241 GAGAAATTTGAGCAGTGGAGCATTTTCTGCTTTGCGAGGTTTCAAAGTGGAGAG 300

QY 1017 ACAACCTATTCTTGAGGGATTTGCAATTTGATCCAAATGTTACTCTGTATGGACTGC 1076  
DB 301 ACAACCTATTCTTGAGGGATTTGCAATTTGATCCAAATGTTACTCTGTATGGACTGC 360

QY 1077 TTCAGGACAGTGTTCATAAAATCATCGTTACAGAGTGCATCTTCTACTGGAGAGGG 1136  
DB 361 TTCAGGACAGTGTTCATAAAATCATCGTTACAGAGTGCATCTTCTACTGGAGAGGG 420

QY 1137 TTCTGTGACTGTGGAGACACAGAGCATGGAACCTGGCCCTTTTGTGTAAATCATGAA 1196  
DB 421 TTCTGTGACTGTGGAGACACAGAGCATGGAACCTGGCCCTTTTGTGTAAATCATGAA 480

QY 1197 CCTGGAAGCAGGATTAATAAGAGAAATTCACGCTGCTCGGTGAAATGAAAGAGGTAAAT 1256  
DB 481 CCTGGAAGCAGGATTAATAAGAGAAATTCACGCTGCTCGGTGAAATGAAAGAGGTAAAT 540

QY 1257 GTCCAGCCAGGAAATATTTCTTCTAGTATATAAATATGCTGTAGAAATGACTATATGG 1316  
DB 541 GTCCAGCCAGGAAATATTTCTTCTAGTATATAAATATGCTGTAGAAATGACTATATGG 600

QY 1317 GAAGGAGAAAGAACTCCCTCTGAACTCCAGATAAGGGAGAAATGAAAGATACTAT 1376  
DB 601 GAAGGAGAAAGAACTCCCTCTGAACTCCAGATAAGGGAGAAATGAAAGATACTAT 660

QY 1377 TGTGTCTCTTTTCAATGATGAACACACATTTCAATATGACCAAGTCAATATACAGCCTACAAAGA 1436  
DB 661 TGTGTCTCTTTTCAATGATGAACACACATTTCAATATGACCAAGTCAATATACAGCCTACAAAGA 720

QY 1437 GCTCTTGTGACTGTGAGCTCCGAGAGCCAGTTGTCATACACATGCTCCATTTGACAAAGAGGT 1496  
DB 721 GCTCTTGTGACTGTGAGCTCCGAGAGCCAGTTGTCATACACATGCTCCATTTGACAAAGAGGT 780

QY 1497 CGTGGGCTGTTAAAGCGGAGCTTATGCTGCTGCCAGGAGCAAGCAAGATATAAAG 1556  
DB 781 CGTGGGCTGTTAAAGCGGAGCTTATGCTGCTGCCAGGAGCAAGCAAGATATAAAG 840

QY 1557 AGTCATTGAGAAATGCTCTCAACATCCACTTCATGTAGAGTATATACACTCAGAGATT 1616  
DB 841 AGTCATTGAGAAATGCTCTCAACATCCACTTCATGTAGAGTATATACACTCAGAGATT 900

QY 1617 ATGGCTCATCAGAAATTTGCTTTGGCTCTTGGTCTCTGATGCAACAAATATATGACTAT 1676  
DB 901 ATGGCTCATCAGAAATTTGCTTTGGCTCTTGGTCTCTGATGCAACAAATATATGACTAT 960

QY 1677 TCAAGTGACTTTAGGCAGATCTTTTGGCAAGCATGCTTTAGAGAAAGACCTGACTCGGAG 1736  
DB 961 TCAAGTGACTTTAGGCAGATCTTTTGGCAAGCATGCTTTAGAGAAAGACCTGACTCGGAG 1020

QY 1737 AATCCCTGCTCATAGCAGAGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTCCCGT 1796  
DB 1021 AATCCCTGCTCATAGCAGAGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTCCCGT 1080

QY 1797 AAGATCCCTCATGATGATCTTCCAGAGTTTATGAGAGTGGATACAAAAATCTC 1856  
DB 1081 AAGATCCCTCATGATGATCTTCCAGAGTTTATGAGAGTGGATACAAAAATCTC 1140

QY 1857 TTTGCTATGGAATTTGTGAAGTATTTATAACAACTGCAAGAAAGATATATCAGTGATGAT 1916  
DB 1141 TTTGCTATGGAATTTGTGAAGTATTTATAACAACTGCAAGAAAGATATATCAGTGATGAT 1200

QY 1917 CATGACAGAGTATCTCTATAACTGCACTTTTCAAGTTCAGATGTTTACTGTTCTTACTCTG 1976  
DB 1201 CATGACAGAGTATCTCTATAACTGCACTTTTCAAGTTCAGATGTTTACTGTTCTTACTCTG 1260

QY 1977 GCTCGACATCTTATTTGAGAGCAGAAATGTTATCTCTGTCATTTACTGAACTCTGTAGAA 2036  
DB 1261 GCTCGACATCTTATTTGAGAGCAGAAATGTTATCTCTGTCATTTACTGAACTCTGTAGAA 1320

QY 2037 GTTTTACCTGAGTACTTTGGAAGAGCAATATAAATTTCAACTTTCCAGGGTTATAGCCAGAC 2096  
DB 1321 GTTTTACCTGAGTACTTTGGAAGAGCAATATAAATTTCAACTTTCCAGGGTTATAGCCAGAC 1380

QY 2097 AAATTTGGAGAGTATATGCACTTATGTCCTTAAAGTATATCTCTGATCAGCAAAACC 2156  
DB 1381 AAATTTGGAGAGTATATGCACTTATGTCCTTAAAGTATATCTCTGATCAGCAAAACC 1440

QY 2157 ACAATATGGAAGAGTATTAAGAAATGCAAGTTCCTTGAAGGTTTTCGATCTTTTGAAG 2216  
DB 1441 ACAATATGGAAGAGTATTAAGAAATGCAAGTTCCTTGAAGGTTTTCGATCTTTTGAAG 1500

QY 2217 ATTCTTACTCTATGTCAGGAAATGCAAGAAATCCGAGACAGGTTGGGCAACACATTCGA 2276  
DB 1501 ATTCTTACTCTATGTCAGGAAATGCAAGAAATCCGAGACAGGTTGGGCAACACATTCGA 1560

QY 2277 GTGGATCTGANTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGATTTTACTC 2336  
DB 1561 GTGGATCTGANTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGATTTTACTC 1620

QY 2337 ATGTTCCAGAGTGGTGTGCTTGTGATGAGAACTCTTACTTGTGGCTTATTAAGAAATGT 2396  
DB 1621 ATGTTCCAGAGTGGTGTGCTTGTGATGAGAACTCTTACTTGTGGCTTATTAAGAAATGT 1680

QY 2397 CACAAAGCTGTGAGGTGAGTACCAAGTTTCTATCTAGTAGCAACAGTAGTAGTACAA 2456  
DB 1681 CACAAAGCTGTGAGGTGAGTACCAAGTTTCTATCTAGTAGCAACAGTAGTAGTACAA 1740

QY 2457 TCGTGTGACATAGTTTGGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTGTGAAGCATA 2516



QY 4677 ACTATCCAGGCAATTGAAATCTATTGGGAGATGAAGGAAACCTCTGTGTTGGAGCACTT 4736  
DB |||||  
QY 3961 ACTATCCAGGCAATTGAAATCTATTGGGAGATGAAGGAAACCTCTGTGTTGGAGCACTT 4020  
DB |||||  
QY 4737 CAAATATAGCAGCATTAATGGTCTGAAAGCATTAATGTCAGTTTGGAGTTGCCAGAGGATT 4796  
DB |||||  
QY 4021 CAAATATAGCAGCATTAATGGTCTGAAAGCATTAATGTCAGTTTGGAGTTGCCAGAGGATT 4080  
DB |||||  
QY 4797 ACCTGTCTCAGGTCCTGATACAGAAACATCTGGTTCGTCTTATCAGTTGTTCTTCTCT 4856  
DB |||||  
QY 4081 ACCTGTCTCAGGTCCTGATACAGAAACATCTGGTTCGTCTTATCAGTTGTTCTTCTCT 4140  
DB |||||  
QY 4857 AACATAAATCAGAGATACACCATGCTTCTGTCTATAGATCTGTTCATAGTTTGGTG 4916  
DB |||||  
QY 4141 AACATAAATCAGAGATACACCATGCTTCTGTCTATAGATCTGTTCATAGTTTGGTG 4200  
DB |||||  
QY 4917 GGTGCTGTGTAGCATCCCATCTTGTATTTGGGATGACCCCTGTGTGATCTGCAGCTTCT 4976  
DB |||||  
QY 4201 GGTGCTGTGTAGCATCCCATCTTGTATTTGGGATGACCCCTGTGTGATCTGCAGCTTCT 4260  
DB |||||  
QY 4977 TCAGTTAGTCTTCTTATAACACCATCTTATCTCTTCCATTTGATCACCATGGCACACATG 5036  
DB |||||  
QY 4261 TCAGTTAGTCTTCTTATAACACCATCTTATCTCTTCCATTTGATCACCATGGCACACATG 4320  
DB |||||  
QY 5037 CTTCAAGATACCTTTACAGTAGACACAGGCTTACCCCTTGTCTCAGGTTCAAGAAGACAGT 5096  
DB |||||  
QY 4321 CTTCAAGATACCTTTACAGTAGACACAGGCTTACCCCTTGTCTCAGGTTCAAGAAGACAGT 4380  
DB |||||  
QY 5097 GAAGAGGCTCATTCGGCATCTTCTTCTTGTGAGAAATTTCTCAATATACAGTGGCTCC 5156  
DB |||||  
QY 4381 GAAGAGGCTCATTCGGCATCTTCTTCTTGTGAGAAATTTCTCAATATACAGTGGCTCC 4440  
DB |||||  
QY 5157 ATTGGGTGTGATATTCCTGGCTGTGATTTGTGGGTCTCACTGAAGAAATGGCATCAACCT 5216  
DB |||||  
QY 4441 ATTGGGTGTGATATTCCTGGCTGTGATTTGTGGGTCTCACTGAAGAAATGGCATCAACCT 4500  
DB |||||  
QY 5217 TATCTTCGCTGTGCTGCTATTTTTCATTTTCACTTTTCTTGGGTAACTCCGCTCAGGAA 5276  
DB |||||  
QY 4501 TATCTTCGCTGTGCTGCTATTTTTCATTTTCTTGGGTAACTCCGCTCAGGAA 4560  
DB |||||  
QY 5277 CTGCATACCAATCTGCAGAGGAGATACAGTGCACTCTGTAGCTATCTATCTTTACCT 5336  
DB |||||  
QY 4561 CTGCATACCAATCTGCAGAGGAGATACAGTGCACTCTGTAGCTATCTATCTTTACCT 4620  
DB |||||  
QY 5337 ACAAAATTTGTTCTCTCTTCCAGCAATATTGGGATACGTAAAGCCCTTGTCTCCAGAGG 5396  
DB |||||  
QY 4621 ACAAAATTTGTTCTCTCTCTTCCAGCAATATTGGGATACGTAAAGCCCTTGTCTCCAGAGG 4680  
DB |||||  
QY 5397 TGGTGTGAGATCTCTGCTTACTAACTGTTTGAAGCAAAAAACACCGTGTGTCAGGTAC 5456  
DB |||||  
QY 4681 TGGTGTGAGATCTCTGCTTACTAACTGTTTGAAGCAAAAAACACCGTGTGTCAGGTAC 4740  
DB |||||  
QY 5457 CTTAGAAAAAGAAATAGTTTGTAGAGTCTCTGATGATATAGTGTGCTCTCTGATCAAA 5516  
DB |||||  
QY 4741 CCTAGAAAAAGAAATAGTTTGTAGAGTCTCTGATGATATAGTGTGCTCTCTGATCAAA 4800  
DB |||||  
QY 5517 GCTTCTCATTTTCAGGTGCCACGGTCTCAGATGATGAGCGAAGCATCTCTCTCTGCTG 5576  
DB |||||  
QY 4801 GCTTCTCATTTTCAGGTGCCACGGTCTCAGATGATGAGCGAAGCATCTCTCTCTGCTG 4860  
DB |||||  
QY 5577 CTTTCTCTGCGGCTATATCTATGTTCTCAGAACATTTGTCGACGAAATTTGTGAACGGG 5636  
DB |||||  
QY 4861 CTTTCTCTGCGGCTATATCTATGTTCTCAGAACATTTGTCGACGAAATTTGTGAACGGG 4920  
DB |||||  
QY 5637 GAAGAGTTGGAGCTTGCAATTTTTCACGCACTTCACTGTGGAGCCGAGTCTGCAATTTTC 5696  
DB |||||  
QY 4921 GAAGAGTTGGAGCTTGCAATTTTTCACGCACTTCACTGTGGAGCCGAGTCTGCAATTTTC 4980  
DB |||||  
QY 5697 CTTAAAAATCAGAGATGCGAGTGTCTGTGTTGAGGTAAAGCCAGAGGCTGTGCTAT 5756  
DB |||||  
QY 4981 CTTAAAAATCAGAGATGCGAGTGTCTGTGTTGAGGTAAAGCCAGAGGCTGTGCTAT 5040  
DB |||||

QY 5757 CCAGTCTTACTTGGATGAATATGGGAAACAGACCCCTGGCTGAAGGGGCAACCCC 5816  
DB |||||  
QY 5041 CCAGTCTTACTTGGATGAATATGGGAAACAGACCCCTGGCTGAAGGGGCAACCCC 5100  
DB |||||  
QY 5817 CTTCAATTTATCTCTGAGCGGTATCGGAAGCTTCCATTTGGTCTGGCAACAACACTGCAATT 5876  
DB |||||  
QY 5101 CTTCAATTTATCTCTGAGCGGTATCGGAAGCTTCCATTTGGTCTGGCAACAACACTGCAATT 5160  
DB |||||  
QY 5877 ATAGAAGAGATTTGTAGGAGCCAAAGAGACTAATCAGATGTTTATTTGGATTTCAACTGGCAG 5936  
DB |||||  
QY 5161 ATAGAAGAGATTTGTAGGAGCCAAAGAGACTAATCAGATGTTTATTTGGATTTCAACTGGCAG 5220  
DB |||||  
QY 5937 TTAATGTGAGCTCCAACTCTGCTCAAGACAAATCAAAATGACGACAGTAAAGGCTG 5996  
DB |||||  
QY 5221 TTAATGTGAGCTCCAACTCTGCTCAAGACAAATCAAAATGACGACAGTAAAGGCTG 5280  
DB |||||  
QY 5997 ATTCAAAATTTATGAAATCTTCTGAGGCTGGGAAAGTATTTGGAGGCTCTTTTCTCCA 6056  
DB |||||  
QY 5281 ATTCAAAATTTATGAAATCTTCTGAGGCTGGGAAAGTATTTGGAGGCTCTTTTCTCCA 5340  
DB |||||  
QY 6057 TGTCCAGGTTCACTTACATCAATATAAATATTTCTTAATGGAGTATTTGCTTTCAATTAGCA 6116  
DB |||||  
QY 5341 TGTCCAGGTTCACTTACATCAATATAAATATTTCTTAATGGAGTATTTGCTTTCAATTAGCA 5400  
DB |||||  
QY 6117 AACATATGCTTCACAGGAAAAAGACATAGATCAATCTGTTTATGCTAGTATTTTCC 6176  
DB |||||  
QY 5401 AACATATGCTTCACAGGAAAAAGACATAGATCAATCTGTTTATGCTAGTATTTTCC 5460  
DB |||||  
QY 6177 AGGAATTTATTTCCCTTCAATATTTGCTCAATTTTATTTTATTTTCCATCCACTTGTAGAT 6236  
DB |||||  
QY 5461 AGGAATTTATTTCCCTTCAATATTTGCTCAATTTTATTTTATTTTCCATCCACTTGTAGAT 5520  
DB |||||  
QY 6237 GAAGTCACTCAACAGTTGTAGACATTTTATGTTGTTAACTCTTCTGCAATTTTGT 6296  
DB |||||  
QY 5521 GAAGTCACTCAACAGTTGTAGACATTTTATGTTGTTAACTCTTCTGCAATTTTGT 5580  
DB |||||  
QY 6297 ATTTGGTGTGTTT 6308  
DB |||||  
QY 5581 ATTTGGTGTGTTT 5592  
DB |||||  
RESULT 4  
ADI16377  
ID ADI16377 standard; DNA; 7648 BP.  
XX AC ADI16377;  
XX AC  
XX AC  
XX 22-APR-2004 (first entry)  
XX Human protein modification and maintenance molecule (PMM) gene #10.  
DE human; protein modification and maintenance molecule; PMM;  
KW gastrointestinal disorders; peptic ulcer; Crohn's disease;  
KW cardiovascular disorders; hypertension; congenital heart disease;  
KW autoimmune disease; inflammatory disease; AIDS; anemia;  
KW developmental disorder; Cushing's syndrome; tubular acidosis;  
KW epithelial disorder; eczema; scabies; neurological disorder;  
KW Alzheimer's disease; multiple sclerosis; infection; cancer; gene; da.  
XX Homo sapiens.  
OS  
XX WO2003100016-A2.  
PN  
XX  
XX  
PD 04-DEC-2003.  
XX  
XX 22-MAY-2003; 2003WO-US016498.  
PF  
XX 22-MAY-2002; 2002US-0383491P.  
PR 24-JUN-2002; 2002US-0391378P.  
PR 22-JUL-2002; 2002US-0397921P.  
XX  
XX (INCY-) INCYTE CORP.  
XX



PI Khare R, Bulloch SA, Swarnakar A, Elliott VS, Marquis JP;  
PI Mason PM, Chawla NK, Ramkumar J, Kabile AE, Hafalia AJA, Lee SY;  
PI Tran UK, Yue H, Becha SD, Griffin JA, Chang H, Jiang X, Jackson AA;  
PI Richardson TW, Lal PG, Yao MG, Lu Y, Warren BA, Jin P, Wilson AD;  
PI Gietzen KU;

XX WPI; 2004-035124/03.

DR P-PSDB; ADI16325.

XX New protein modification and maintenance molecules, useful for diagnosing  
PT or treating e.g. peptic ulcer, hypertension, rheumatic fever, AIDS,  
PT Cushing's syndrome, Alzheimer's disease, multiple sclerosis, stroke or  
PT cancers.

XX Claim 5; SEQ ID NO 62; 419pp; English.

XX The invention comprises the amino acid and coding sequences of human  
CC protein modification and maintenance molecules (PMM). The DNA and  
CC protein sequences of the invention are useful for the diagnosis and  
CC treatment of disorders associated with expression of PMM, such as:  
CC gastrointestinal disorders (e.g. peptic ulcer and Crohn's disease),  
CC cardiovascular disorders (e.g. hypertension and congenital heart  
CC disease), autoimmune or inflammatory disease (e.g. AIDS and anaemia),  
CC developmental disorders (e.g. Cushing's syndrome and tubular acidosis),  
CC epithelial disorders (e.g. eczema and scabies), neurological disorders  
CC (e.g. Alzheimer's disease and multiple sclerosis), infections and cancer.  
XX The present DNA sequence encodes a human PMM protein of the invention.

SQ Sequence 7648 BP; 2355 A; 1407 C; 1606 G; 2280 T; 0 U; 0 Other;

Query Match 85.4%; Score 5388; DB 12; Length 7648;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 5498; Conservative 0; Mismatches 0; Indels 100; Gaps 1;  
  
QY 711 GCTGAGGTACTGACAGGATGAAATCAGCGGGAGTTACCCAGACCCCTCAGGCTGTG 770  
DB 1 GCTGAGGTACTGACAGGATGAAATCAGCGGGAGTTACCCAGACCCCTCAGGCTGTG 60  
  
QY 771 GCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTGCATCATTGGCACAA 830  
DB 61 GCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTGCATCATTGGCACAA 120  
  
QY 831 TTGGTGCCAGAAATTTACTTTCTCAATGGACCCAGACTTGGAAAGCAGGAGGAAAGT 890  
DB 121 TTGGTGCCAGAAATTTACTTTCTGCTGAATGGACCCAGACTTGGAAAGCAGGAGGAAAGT 180  
  
QY 891 GTACAAATGTCAATATTCATCCACTCGAATGGTACTTATTTGGAGAGATCCAGATATT 950  
DB 181 GTACAAATGTCAATATTCATCCACTCGAATGGTACTTATTTGGAGAGATCCAGATATT 240  
  
QY 951 TGCTTAGAGAAATTCAGACAGTGGAGCATTTTCAGCTTTGGGAGGGTTTCAAAAGT 1010  
DB 241 TGCTTAGAGAAATTCAGACAGTGGAGCATTTTCAGCTTTGGGAGGGTTTCAAAAGT 300  
  
QY 1011 GGAGAGACAACTATTCTTGGAGGATTTGTGCAATTTGATCCAAATGTGTTACTCTGTATG 1070  
DB 301 GGAGAGACAACTATTCTTGGAGGATTTGTGCAATTTGATCCAAATGTGTTACTCTGTATG 360  
  
QY 1071 GACTGCTTCCAGGACAGTGTTCATAAAAATCAATCGTTTCAAGATGCATCTTCTACTGGA 1130  
DB 361 GACTGCTTCCAGGACAGTGTTCATAAAAATCAATCGTTTCAAGATGCATCTTCTACTGGA 420  
  
QY 1131 GGAGGGTTCTGTGACTGTGGAGACACAGAGGCATCGAAATCTGGGCCCTTTTGTGTAAAT 1190  
DB 421 GGAGGGTTCTGTGACTGTGGAGACACAGAGGCATCGAAATCTGGGCCCTTTTGTGTAAAT 480  
  
QY 1191 CATGAACCTGGAAGACAGGTACTATATAAGAGAAATTCAGCTGTCCGTTTGAATCAAGAG 1250  
DB 481 CATGAACCTGGAAGACAGGTACTATATAAGAGAAATTCAGCTGTCCGTTTGAATCAAGAG 540  
  
QY 1251 GTAATTGTCCAGCCAGGAAATATTTCTTCCAGTGATAAAAATATGTGCTAGAAATGACT 1310  
DB 541 GTAATTGTCCAGCCAGGAAATATTTCTTCCAGTGATAAAAATATGTGCTAGAAATGACT 600

QY 1311 ATATGGGAAGAGGAAAAAGAACTGCCTCCTGAACTCCAGATTAAGGAGAAAAATGAAGA 1370  
DB 601 ATATGGGAAGAGGAAAAAGAACTGCCTCCTGAACTCCAGATTAAGGAGAAAAATGAAGA 660  
  
QY 1371 TACTATTGTGTCTTTTCAATGATGAACACCAATTCATATGACCAAGCTCATATACAGCTA 1430  
DB 661 TACTATTGTGTCTTTTCAATGATGAACACCAATTCATATGACCAAGCTCATATACAGCTA 720  
  
QY 1431 CAAGAGCTCTTGACTGTGAGCTGCGAGAGCCAGTTGCATACATCCATGCTCCATTCAGAAA 1490  
DB 721 CAAGAGCTCTTGACTGTGAGCTGCGAGAGCCAGTTGCATACATCCATGCTCCATTCAGAAA 780  
  
QY 1491 GAGGCTCGTGGGCTGTTAAAGCGGGAGCTTATCTGCTTGCAGGAGCAAGCAAGGAAGAT 1550  
DB 781 GAGGCTCGTGGGCTGTTAAAGCGGGAGCTTATCTGCTTGCAGGAGCAAGCAAGGAAGAT 840  
  
QY 1551 ATAAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACTTCATGTAGAGTATTTACACTCA 1610  
DB 841 ATAAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACTTCATGTAGAGTATTTACACTCA 900  
  
QY 1611 GAGATTATGGCTCATCAGAAATTTGCTTGGCTCTTGGTCTTGGATGCAACAAATATTATG 1670  
DB 901 GAGATTATGGCTCATCAGAAATTTGCTTGGCTCTTGGTCTTGGATGCAACAAATATTATG 960  
  
QY 1671 AGCTATTCAAGTACTTTTAGGCAGATCTTTTGGCCAAAGCATGCTTTAGAGAAAGAACCTGAC 1730  
DB 961 AGCTATTCAAGTACTTTTAGGCAGATCTTTTGGCCAAAGCATGCTTTAGAGAAAGAACCTGAC 1020  
  
QY 1731 TCGAGAGATCCCTGTCTCTATAAGCAGGTTAATGTTGGGATGCAAGCTTTTATAAGGT 1790  
DB 1021 TCGAGAGATCCCTGTCTCTATAAGCAGGTTAATGTTGGGATGCAAGCTTTTATAAGGT 1080  
  
QY 1791 GCCCGTAGAGTCTTTCATGAATTCATCTTCAGCAGTTTATTTTATGGAGATGGAATACAAA 1850  
DB 1081 GCCCGTAGAGTCTTTCATGAATTCATCTTCAGCAGTTTATTTTATGGAGATGGAATACAAA 1140  
  
QY 1851 AAACCTCTTTGCTATGGAATTTGTAAGTATTATAAACAATCTGCAGAAAGAAATATATCAGT 1910  
DB 1141 AAACCTCTTTGCTATGGAATTTGTAAGTATTATAAACAATCTGCAGAAAGAAATATATCAGT 1200  
  
QY 1911 GATGATCATGACAGAGTATCTTATAACTGCACTTTTCAGTTTTCAGTTTTCAGTTTTCCT 1970  
DB 1201 GATGATCATGACAGAGTATCTTATAACTGCACTTTTCAGTTTTCAGTTTTCAGTTTTCCT 1260  
  
QY 1971 ACTCTGGCTCGACATCTTTTGAAGCAGAGTATCTCTCTGTCATTTACCTGAACTCTG 2030  
DB 1261 ACTCTGGCTCGACATCTTTTGAAGCAGAGTATCTCTCTGTCATTTACCTGAACTCTG 1320  
  
QY 2031 CTAGAAGTTTTCCTGAGTACTTGGACAGGAAACAATAAAATTCACACTTCACGGGTTATAGC 2090  
DB 1321 CTAGAAGTTTTCCTGAGTACTTGGACAGGAAACAATAAAATTCACACTTCACGGGTTATAGC 1380  
  
QY 2091 CAGACAAATTTGGGAAGATATATGCGATATATGTCACCTTAAAGTATATCTGATCAGC 2150  
DB 1381 CAGACAAATTTGGGAAGATATATGCGATATATGTCACCTTAAAGTATATCTGATCAGC 1440  
  
QY 2151 AAACCCCAATATGGAAGAGATTAAGAAATGCAATTTTGAAGGTTTTCGATCTTTT 2210  
DB 1441 AAACCCCAATATGGAAGAGATTAAGAAATGCAATTTTGAAGGTTTTCGATCTTTT 1500  
  
QY 2211 TTGAAGATTTCTTACTGATGAGAGATTAAGAAATGCAATTTTGAAGGTTTTCGATCTTTT 2270  
DB 1501 TTGAAGATTTCTTACTGATGAGAGATTAAGAAATGCAATTTTGAAGGTTTTCGATCTTTT 1560  
  
QY 2271 ATTGAAGTGGATCTGATTTGGAGGCTGCCATTTGCTATATACAGATGCAATTTGAAGATATT 2330  
DB 1561 ATTGAAGTGGATCTGATTTGGAGGCTGCCATTTGCTATATACAGATGCAATTTGAAGATATT 1620  
  
QY 2331 TTACTGATTTCCAGAGTGGTGTGTCATGAGAACTCTTACTTGTGCTTATATAA 2390  
DB 1621 TTACTGATTTCCAGAGTGGTGTGTCATGAGAACTCTTACTTGTGCTTATATAA 1680

Qy	2391	GAATGTCACAAAGCTGTGATGAGGTGCAGTACCAAGTTTCATATCTAGTAGCAAGACAGTA	2455
Db	1681	GAATGTCACAAAGCTGTGATGAGGTGCAGTACCAAGTTTCATATCTAGTAGCAAGACAGTA	1740
Qy	2451	GTACAAATCGTGTGCACATAGTTTGGAAACAAAGTCCTACAGAGTATCTGAGGATCTTGTA	2510
Db	1741	GTACAAATCGTGTGCACATAGTTTGGAAACAAAGTCCTACAGAGTATCTGAGGATCTTGTA	1800
Qy	2511	AGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTCATGTACGTTTAAAGCAGGCTG	2570
Db	1801	AGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTCATGTACGTTTAAAGCAGGCTG	1860
Qy	2571	GGTCTGTTTCAAGACTGCATGAATTTGTGTCTTTTGAGGACTTTCAAGTAGAGTACTA	2630
Db	1861	GGTCTGTTTCAAGACTGCATGAATTTGTGTCTTTTGAGGACTTTCAAGTAGAGTACTA	1920
Qy	2631	GTGGAAATATCCTTTACGTTGCTCGTGTGTGGTTCGCCAGGTTGTGCTGAGATGTGGCGA	2690
Db	1921	GTGGAAATATCCTTTACGTTGCTCGTGTGTGGTTCGCCAGGTTGTGCTGAGATGTGGCGA	1980
Qy	2691	AGAAATGGACTGTCTCTTATTAGCCAGGTGTTTTATTACCAAGATGTTAAAGTCAGAGAA	2750
Db	1981	AGAAATGGACTGTCTCTTATTAGCCAGGTGTTTTATTACCAAGATGTTAAAGTCAGAGAA	2040
Qy	2751	GAATGTATGATTAAGATATCATCATGCTTCAGATGGTGATCTTTAATGTGATCCCAAT	2810
Db	2041	GAATGTATGATTAAGATATCATCATGCTTCAGATGGTGATCTTTAATGTGATCCCAAT	2100
Qy	2811	AAGTTCCTGTGTACCTGTTACTTTCAGAGTATGAACTTGCOCGAGGCTTTTACACAGACCAT	2870
Db	2101	AAGTTCCTGTGTGTACCTGTTACTTTCAGAGTATGAACTTGCOCGAGGCTTTTACACAGACCAT	2160
Qy	2871	TCTACAAAAGACACAGGATTTGATTTAAACAATATAATACATCTAATAGAGAAATGCTTCAG	2930
Db	2161	TCTACAAAAGACACAGGATTTGATTTAAACAATATAATACATCTAATAGAGAAATGCTTCAG	2220
Qy	2931	GTCCTCATCTATATGTGGGTGAGCGTTATGTACTGGAGTGGGAAATGTGACCAAGAA	2990
Db	2221	GTCCTCATCTATATGTGGGTGAGCGTTATGTACTGGAGTGGGAAATGTGACCAAGAA	2280
Qy	2991	GAGGTCAATGAGAGAAATCATCTGCTTTGCAATTGAACCCATGCCACACAGTGCC	3050
Db	2281	GAGGTCAATGAGAGAAATCATCTGCTTTGCAATTGAACCCATGCCACACAGTGCC	2340
Qy	3051	ATTGCCAAAATTTTACTCGAGATGAAAATAATGAAACTGGCTTAGAGAAATGCTATAAAC	3110
Db	2341	ATTGCCAAAATTTTACTCGAGATGAAAATAATGAAACTGGCTTAGAGAAATGCTATAAAC	2400
Qy	3111	AAAGTGGCCACATTTAAGAAACCGAGTGTATCAGGCCATGGAGTTTATGAATTAAGAT	3170
Db	2401	AAAGTGGCCACATTTAAGAAACCGAGTGTATCAGGCCATGGAGTTTATGAATTAAGAT	2460
Qy	3171	GAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAAACCCAGCATAGCAAG	3230
Db	2461	GAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAAACCCAGCATAGCAAG	2520
Qy	3231	GCTGAAACATATGCAGAGAAAGAGAGAAAACAAGAAAACAAAGATGAAGCATTTGCCGCCA	3290
Db	2521	GCTGAAACATATGCAGAGAAAGAGAGAAAACAAGAAAACAAAGATGAAGCATTTGCCGCCA	2580
Qy	3291	CCACCCTCCTGAAATTCGCCCTGCTTTTCAGCAAAGTGAATTAACCTTCTCAACTGTGAT	3350
Db	2581	CCACCCTCCTGAAATTCGCCCTGCTTTTCAGCAAAGTGAATTAACCTTCTCAACTGTGAT	2640
Qy	3351	ATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCTAACTTG	3410
Db	2641	ATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCTAACTTG	2700
Qy	3411	TGGACCGAAGGATGTCTCCAAAATGGCTTTTTCATATTTCTGGCAATTTGGGTTTCTAGAAAGAG	3470
Db	2701	TGGACCGAAGGATGTCTCCAAAATGGCTTTTTCATATTTCTGGCAATTTGGGTTTCTAGAAAGAG	2724
Qy	3471	AAGCAACAGCTTTCAAAAAGCTCTCTGAAAGAGAAAGTAACTTTTGACTTTTATCATAAGCT	3530

Db	2725	-----	2726
Qy	3531	TCAGATTGGGAAGTTCCAGCCATGAATATACAAATGCTTTTGGAAAAAATCAAGGAAT	3590
Db	2725	-----GATTGGGAAGTTCCAGCCATGAATATACAAATGCTTTTGGAAAAAATCAAGGAAT	2780
Qy	3591	CCCCAGTTAGAAAGCCAGAGGACATGATAAACGTGGATACTTCAGATGTTTGGACACAGTG	3650
Db	2781	CCCCAGTTAGAAAGCCAGAGGACATGATAAACGTGGATACTTCAGATGTTTGGACACAGTG	2840
Qy	3651	AAGCGATTAGAGAGAAAAATCTTGTTTAAATGTAGCAACACACATCAGGATCGGAATCTATT	3710
Db	2841	AAGCGATTAGAGAGAAAAATCTTGTTTAAATGTAGCAACACACATCAGGATCGGAATCTATT	2900
Qy	3711	AAGATGATCAGATTAATCTATGATAAAGAAAAAGCAGAACGAAAAAGCTGAAGCT	3770
Db	2901	AAGATGATCAGATTAATCTATGATAAAGAAAAAGCAGAACGAAAAAGCTGAAGCT	2960
Qy	3771	GCTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGCTTGCCTTTACAGAAAACTTCATT	3830
Db	2961	GCTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGCTTGCCTTTACAGAAAACTTCATT	3020
Qy	3831	GAAACTCATAACTCATGTATGATCAATACATCAGAAATGCTCGGGAAGAAAGATTCATT	3890
Db	3021	GAAACTCATAACTCATGTATGATCAATACATCAGAAATGCTCGGGAAGAAAGATTCATT	3080
Qy	3891	ATGAGAGAGAGAGACCCCCAGCAGTCAGTGAATCTCTAGAAATGCTTTGGTCTCTAAA	3950
Db	3081	ATGAGAGAGAGAGACCCCCAGCAGTCAGTGAATCTCTAGAAATGCTTTGGTCTCTAAA	3140
Qy	3951	CGGGTCCATCTGTACTGAAAAGGAGGTGCTCAGCTGCATCCTTTGCCAAGAAAGAACAG	4010
Db	3141	CGGGTCCATCTGTACTGAAAAGGAGGTGCTCAGCTGCATCCTTTGCCAAGAAAGAACAG	3200
Qy	4011	GAGGTGAAAATAGAAAATATGCCATGGTATTAATCGGCCCTGTGTCCAGAAATCTACTGCC	4070
Db	3201	GAGGTGAAAATAGAAAATATGCCATGGTATTAATCGGCCCTGTGTCCAGAAATCTACTGCC	3260
Qy	4071	TTAAACCCAGACAGGGGAAAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTC	4130
Db	3261	TTAAACCCAGACAGGGGAAAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTC	3320
Qy	4131	ATGGATCCAGACTTGGCATATGAACTTATATACAGGAAGCTGTGGTCACTTAATGACAGCA	4190
Db	3321	ATGGATCCAGACTTGGCATATGAACTTATATACAGGAAGCTGTGGTCACTTAATGACAGCA	3380
Qy	4191	GTGTGCTGGCAGAAATATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTT	4250
Db	3381	GTGTGCTGGCAGAAATATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTT	3440
Qy	4251	GACCTTTTTCAGCTTGGAAAGTGAGAAATATCTTTGGCCCTCTTTGCAAAATCTCTGTGCAAT	4310
Db	3441	GACCTTTTTCAGCTTGGAAAGTGAGAAATATCTTTGGCCCTCTTTGCAAAATCTCTGTGCAAT	3500
Qy	4311	ACTGTGATCCCAATATTTCTTTGCAACTCAAAGATAAACAGTGAGAAATGACAGATGCT	4370
Db	3501	ACTGTGATCCCAATATTTCTTTGCAACTCAAAGATAAACAGTGAGAAATGACAGATGCT	3560
Qy	4371	CTTGCTCAACTTTTGACCCCTGGCACGGTGGATACAGATGTTCTTGCCCAAGAAATATCAGGT	4430
Db	3561	CTTGCTCAACTTTTGACCCCTGGCACGGTGGATACAGATGTTCTTGCCCAAGAAATATCAGGT	3620
Qy	4431	TATAATATAGACATGCTTAAGAGAGAAACCCAAATTTCTTATTTCTTTTAAATCAAGGAATG	4490
Db	3621	TATAATATAGACATGCTTAAGAGAGAAACCCAAATTTCTTATTTCTTTTAAATCAAGGAATG	3680
Qy	4491	GGAGATTTCTACTTTGGAGTTCCATTCATCCTCAGTTTTTGGCGTTGAGTCTTCGATTTAAA	4550
Db	3681	GGAGATTTCTACTTTGGAGTTCCATTCATCCTCAGTTTTTGGCGTTGAGTCTTCGATTTAAA	3740
Qy	4551	TATTCAAATAGCATCAAGGAATATGGTTATTTCTTTTGCCCAACAAATTTATAGAAATGGA	4610

```
Db 3741 TATTCAATATAGCATCAAGAAATGGTTATTCTCTTTGGCCACAACAATTTATAGAAATGGA 3800
Qy 4611 TTGAAAGTGCACCTGATGAAAGGATCTCGAGTCCCATGCTGACCTGGAGCACTGC 4670
Db 3801 TTGAAAGTGCACCTGATGAAAGGATCTCGAGTCCCATGCTGACCTGGAGCACTGC 3860
Qy 4671 GCTTTCACTATCTCAGGCAATTTGMAAATCTATTGGGAGATGAAGGAAACCTCTGTTTGA 4730
Db 3861 GCTTTCACTATCTCAGGCAATTTGMAAATCTATTGGGAGATGAAGGAAACCTCTGTTTGA 3920
Qy 4731 GCATTTCAAAATAGGACGATTAATGGTCTGAAAGCATTAATGCAAGTTTGCAGTTGCACAG 4790
Db 3921 GCATTTCAAAATAGGACGATTAATGGTCTGAAAGCATTAATGCAAGTTTGCAGTTGCACAG 3980
Qy 4791 AGGATTACTGTCTCAGGTCTGATACAGAAACATCTGGTTGTTCTCTATCAGTTGTT 4850
Db 3981 AGGATTACTGTCTCAGGTCTGATACAGAAACATCTGGTTGTTCTCTATCAGTTGTT 4040
Qy 4851 CTTCCTAACATAAAATCAGAAGATACACCATGCTCTCTGTCTATAGATCTGTTTCATGTT 4910
Db 4041 CTTCCTAACATAAAATCAGAAGATACACCATGCTCTCTGTCTATAGATCTGTTTCATGTT 4100
Qy 4911 TTGGTGGGTGCTGTGTAGCATTCCTCATCTTTGATTTGGGATGACCCCTGTTGATCTGCAG 4970
Db 4101 TTGGTGGGTGCTGTGTAGCATTCCTCATCTTTGATTTGGGATGACCCCTGTTGATCTGCAG 4160
Qy 4971 CTTCTTCAGTTAGTCTTCTTAACCACTTTATCTCTTCCATTTGATCACCATTGGCA 5030
Db 4161 CTTCTTCAGTTAGTCTTCTTAACCACTTTATCTCTTCCATTTGATCACCATTGGCA 4220
Qy 5031 CACATGCTTCAGATACTACTTACAGTAGACACAGGCCCTACCCCTTGCTCAGTTCAAGAA 5090
Db 4221 CACATGCTTCAGATACTACTTACAGTAGACACAGGCCCTACCCCTTGCTCAGTTCAAGAA 4280
Qy 5091 GACAGTGAAGAGGCTCATTCGGCATCTTCTTTCTTGAGAAATTTCTCAATATACAGT 5150
Db 4281 GACAGTGAAGAGGCTCATTCGGCATCTTCTTTCTTGAGAAATTTCTCAATATACAGT 4340
Qy 5151 GGCTCCATTTGGTGTGATATCTCGCTGGTGTATTGTTGGGGTCTCATGAGAAATGGCATC 5210
Db 4341 GGCTCCATTTGGTGTGATATCTCGCTGGTGTATTGTTGGGGTCTCATGAGAAATGGCATC 4400
Qy 5211 ACCCTTATCTCGCTGTGCTGATGTTGTTTCCACTATTTACTTTGGGGTAACTCCGCCT 5270
Db 4401 ACCCTTATCTCGCTGTGCTGATGTTGTTTCCACTATTTACTTTGGGGTAACTCCGCCT 4460
Qy 5271 GAGGAACCTGCATACCAATTTCTGCAGAGGAGATACAGTGCACCTCTGTAGCTATCTATCT 5330
Db 4461 GAGGAACCTGCATACCAATTTCTGCAGAGGAGATACAGTGCACCTCTGTAGCTATCTATCT 4520
Qy 5331 TTACTCTACAAATTTGTTCTGCTCTTCCAGGAATATTGGGATACATGTAAGGCCCTTGCTC 5390
Db 4521 TTACTCTACAAATTTGTTCTGCTCTTCCAGGAATATTGGGATACATGTAAGGCCCTTGCTC 4580
Qy 5391 CAGAGTGTGTGCAGATCTCGCTTACTAACTGTTTGAAGCAAAAAACACCGTGTCT 5450
Db 4581 CAGAGTGTGTGCAGATCTCGCTTACTAACTGTTTGAAGCAAAAAACACCGTGTCT 4640
Qy 5451 AGGTACCTTAGAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGACTATAGCTGCTCCTG 5510
Db 4641 AGGTACCTTAGAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGACTATAGCTGCTCCTG 4700
Qy 5511 AATCAAGCTTCTCATTTCAAGTGTGCCAGGTCTGCAGATGATGACGGAAGCATCTGTC 5570
Db 4701 AATCAAGCTTCTCATTTCAAGTGTGCCAGGTCTGCAGATGATGACGGAAGCATCTGTC 4760
Qy 5571 CTCTGCCCTTTCTGTTGGGCTATATCTATGTTCTCAGAACATTTGCTGCGAGGAATTTGT 5630
Db 4761 CTCTGCCCTTTCTGTTGGGCTATATCTATGTTCTCAGAACATTTGCTGCGAGGAATTTGT 4820
Qy 5631 AACGGGGAAGAGTTTGGAGCTTGCATTTTTTTCACGCACTTTCATCTGTGGAGCGGAGTCTGC 5690
Db 4821 AACGGGGAAGAGTTTGGAGCTTGCATTTTTTTCACGCACTTTCATCTGTGGAGCGGAGTCTGC 4880
```

```
Qy 5691 ATTTTCTTAAAAATCAGAGATGCCAGTGTCTGTTGTTGAAGTAAAGCCAGAGGCTGT 5750
Db 4881 ATTTTCTTAAAAATCAGAGATGCCAGTGTCTGTTGTTGAAGTAAAGCCAGAGGCTGT 4940
Qy 5751 GCCTATCCAGCTCCTTACTTGGATGAATATGGAGAAAACAGACCTCGCTGGAAGGGGC 5810
Db 4941 GCCTATCCAGCTCCTTACTTGGATGAATATGGAGAAAACAGACCTCGCTGGAAGGGGC 5000
Qy 5811 AACCCCTTCAATTATCTCTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACAC 5870
Db 5001 AACCCCTTCAATTATCTCTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACAC 5060
Qy 5871 TGCAATTATAGAAGAGATTGCTTAGGAGCAAGAGACTAATCAGATGTTATTTGGATTCAAC 5930
Db 5061 TGCAATTATAGAAGAGATTGCTTAGGAGCAAGAGACTAATCAGATGTTATTTGGATTCAAC 5120
Qy 5931 TGGCAGTTACTGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAA 5990
Db 5121 TGGCAGTTACTGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAA 5180
Qy 5991 AGGCTGATTCAAAATTAAGGAAAACCTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTCTTTT 6050
Db 5181 AGGCTGATTCAAAATTAAGGAAAACCTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTCTTTT 5240
Qy 6051 GCTCCATGTCAGGTTCACTTTACATCAATAAAATATTCTTAAATGGAGTATTGCTTTCAA 6110
Db 5241 GCTCCATGTCAGGTTCACTTTACATCAATAAAATATTCTTAAATGGAGTATTGCTTTCAA 5300
Qy 6111 TTAGCAACATATGCTTTACAGGAAAAAGACATAGATCAATCTGTTTTATATGCTAGT 6170
Db 5301 TTAGCAACATATGCTTTACAGGAAAAAGACATAGATCAATCTGTTTTATGCTAGT 5360
Qy 6171 ATTTCCAGGAATTTATTTCCCTTCATAATTTGCTCATTTCAATTTATTTCACTCCACTG 6230
Db 5361 ATTTCCAGGAATTTATTTCCCTTCATAATTTGCTCATTTCAATTTATTTCACTCCACTG 5420
Qy 6231 GTAGATCAAGTTCAGTCAAAACAGTTGTAGACATTTTATGTTGTTAACTCTTCTGCAA 6290
Db 5421 GTAGATCAAGTTCAGTCAAAACAGTTGTAGACATTTTATGTTGTTAACTCTTCTGCAA 5480
Qy 6291 TTTTGTATTTTGGTGTTTT 6308
Db 5481 TTTTGTATTTTGGTGTTTT 5498
```

## RESULT 5

ADJ95458  
ID ADJ95458 standard; cDNA; 5205 BP.

AC ADJ95458;

XX 03-JUN-2004 (first entry)

XX Human cDNA encoding Ubiquitin ligase E3alpha I, synthetic variant.

XX Human; ss; gene; ubiquitin ligase; E3alpha I;

KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;

KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;

KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;

KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;

KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome;

KW SNP; single nucleotide polymorphism.

XX Homo sapiens.

OS Synthetic.

XX Location/Qualifiers

PH 1..5205

FT /\*tag= a

FT /product= "Amino acids 6-1734 of SEQ ID 19"

FT /transl\_except= (pos:658..678,aa:REKNERYCYVLFNDEHHSHYDHF)

FT /transl\_except= (pos:4933..4998,aa:CK)

variation replace(4657,T)  
/\*tag= b  
/standard\_name= "single nucleotide polymorphism"  
US6706505-B1.  
16-MAR-2004.  
28-NOV-2000; 2000US-007241126.  
08-MAR-2000; 2000US-0187911P.  
(AMGE-) AMGEN INC.  
PI Han H, Kwak K;  
XX WPI; 2004-236723/22.  
DR P-PSDB; ADJ95459.  
XX New nucleic acid molecule, useful for preparing a composition for  
PT diagnosing, treating or preventing diseases associated with human  
PT E3approximately polypeptide, e.g., muscle atrophy.  
XX  
PS Claim 19; SEQ ID NO 18; 104pp; English.  
XX  
CC The invention relates to a new isolated nucleic acid molecule appearing  
CC as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I  
CC protein appearing as ADJ95442. Also included are a vector comprising the  
CC nucleic acid, a host cell comprising the vector, a process of producing a  
CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the  
CC nucleic acid molecule, a reagent comprising a detectably labelled  
CC nucleotide, and a method for determining the presence of a human E3alpha  
CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid  
CC molecule is useful for preparing a composition for diagnosing, treating  
CC or preventing diseases associated with human E3alpha I polypeptide, e.g.  
CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal  
CC cachexia, inflammatory cachexia, muscle wasting disorders associated with  
CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,  
CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting  
CC syndrome. The present sequence encodes a human E3alpha I variant. NOTE:  
CC This cDNA was assembled from previously isolated cDNA fragments and is  
CC stated to encode ADJ95459, but appears to be incomplete as it only  
CC encodes from amino acid 6 and has translation exceptions.  
XX  
SQ Sequence 5205 BP; 1553 A; 1017 C; 1171 G; 1450 T; 0 U; 14 Other;  
Query Match 81.4%; Score 5136; DB 12; Length 5205;  
Best Local Similarity 98.8%; Pred No. 0;  
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;  
QY 696 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTACCCAG 755  
DB 1 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTACCCAG 60  
QY 756 ACCCTCAGCGTCTGGCATCTTGGTGGATGAGCAAGTTGATTTTATCTGCTTTCTTG 815  
DB 61 ACCCTCAGCGTCTGGCATCTTGGTGGATGAGCAAGTTGATTTTATCTGCTTTCTTG 120  
QY 816 CATCATTTGGCACAATTTGGTCCAGAAATTTACTTTGCTGAAATGGAATCCAGACTTGGA 875  
DB 121 CATCATTTGGCACAATTTGGTCCAGAAATTTACTTTGCTGAAATGGAATCCAGACTTGGA 180  
QY 876 AAGCAGGAGGAAAGTGAATGTCATATTTCAATTTCACTCCACTGGAATGGTACTTTTGA 935  
DB 181 AAGCAGGAGGAAAGTGAATGTCATATTTCAATTTCACTCCACTGGAATGGTACTTTTGA 240  
QY 936 GAAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCACTGGAGCAATTTTCACTTTGGG 995  
DB 241 GAAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCACTGGAGCAATTTTCACTTTGGG 300  
QY 996 AGGGTTTTCAAGTGGAGGACACCTATTCTTTCAGGGATTTGCAATTTGATCCACACA 1055  
DB 301 AGGGTTTTCAAGTGGAGGACACCTATTCTTTCAGGGATTTGCAATTTGATCCACACA 360

QY 1056 TGTGTACTCTGTATGGACTGCTTTCAGGACAGTGTTCATATAAATCATCGTTTACAAGATG 1115  
DB 361 TGTGTACTCTGTATGGACTGCTTTCAGGACAGTGTTCATATAAATCATCGTTTACAAGATG 420  
QY 1116 CATACTTCTACTGGAGGAGGGTTCGTGACTGTGTGAGACACAGAGGCAATGGAATACTGGC 1175  
DB 421 CATACTTCTACTGGAGGAGGGTTCGTGACTGTGTGAGACACAGAGGCAATGGAATACTGGC 480  
QY 1176 CCTTTTGTGTAAATCATGAACTCTGAAAGAGCAGGTACTATATAAAGAGAAATTCAGCTGT 1235  
DB 481 CCTTTTGTGTAAATCATGAACTCTGAAAGAGCAGGTACTATATAAAGAGAAATTCAGCTGT 540  
QY 1236 CCGTTGAATGAAGAGGTAAATTTGTCCAAGCCAGGAAATATTTCTTCTCAGTGATAAATAT 1295  
DB 541 CCGTTGAATGAAGAGGTAAATTTGTCCAAGCCAGGAAATATTTCTTCTCAGTGATAAATAT 600  
QY 1296 GTCTGTAGAAATGCACTATATGGAAGAGGAAAGAACTGCTTCTGAACTCCAGATAAGG 1355  
DB 601 GTCTGTAGAAATGCACTATATGGAAGAGGAAAGAACTGCTTCTGAACTCCAGATAAGG 660  
QY 1356 GAGAAAAATGAAAGATACCTATTTGTCTCTTTTCAATGATGAACACCATTCATATGACCA 1415  
DB 661 KNR-----YVCVNDHSHVDH 675  
QY 1416 GTCATATACAGCTTACAAAGAGCTTTGACTGTGTGAGCTCGCAGAGGCCAGTTGCATACC 1475  
DB 676 GTCATATACAGCTTACAAAGAGCTTTGACTGTGTGAGCTCGCAGAGGCCAGTTGCATACC 735  
QY 1476 ACTGCCATTGACAAAGAGGGTCTGGGCTGTGTTAAAGCGGGAGCTTATGCTGCTTGGCCAG 1535  
DB 736 ACTGCCATTGACAAAGAGGGTCTGGGCTGTGTTAAAGCGGGAGCTTATGCTGCTTGGCCAG 795  
QY 1536 GAAGCAAGGAGGATATATAAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 1595  
DB 796 GAAGCAAGGAGGATATATAAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 855  
QY 1596 GAAGTATTTACACTCAGAGATTTATGGCTCATCAGAAAAATTTGCTTTCGCTTGGTTCCTGG 1655  
DB 856 GAAGTATTTACACTCAGAGATTTATGGCTCATCAGAAAAATTTGCTTTCGCTTGGTTCCTGG 915  
QY 1656 ATGAACAAAAATATGAGCTATTTCAAGTCATTTTAGGAGAGATCTTTTCCAGCAGCATGCTT 1715  
DB 916 ATGAACAAAAATATGAGCTATTTCAAGTCATTTTAGGAGAGATCTTTTCCAGCAGCATGCTT 975  
QY 1716 AGAGAGAAACCTGACTCGGAGAACTCCCTGTCTCTATAAGCAGGTTAATGCTTTGGGATGCA 1775  
DB 976 AGAGAGAAACCTGACTCGGAGAACTCCCTGTCTCTATAAGCAGGTTAATGCTTTGGGATGCA 1035  
QY 1776 AAGCTTTATAAAGGTGCCGTAAGATCCTTCATGAATTTGATCTTCAGCAGTTTTTTTATG 1835  
DB 1036 AAGCTTTATAAAGGTGCCGTAAGATCCTTCATGAATTTGATCTTCAGCAGTTTTTTTATG 1095  
QY 1836 GAGATGGATACAAAAAATCTTTTGTCTATGAAATTTGTGAAGTATTAATAACACTGAG 1895  
DB 1096 GAGATGGATACAAAAAATCTTTTGTCTATGAAATTTGTGAAGTATTAATAACACTGAG 1155  
QY 1896 AAAGAAATATACAGTGATGATCATGACAGAAAGTATCTCTATAAGTCACTTTTCACTTCAG 1955  
DB 1156 AAAGAAATATACAGTGATGATCATGACAGAAAGTATCTCTATAAGTCACTTTTCACTTCAG 1215  
QY 1956 ATGTTTACTGTTCTTACTCTGGCTCGACATCTTATTTGAAGAGCAGAAATGTTTATCTCTGC 2015  
DB 1216 ATGTTTACTGTTCTTACTCTGGCTCGACATCTTATTTGAAGAGCAGAAATGTTTATCTCTGC 1275  
QY 2016 ATTACTGAATCTGCTAGAAATTTTACTGAGTACTTTGGACAGGAAATAAATTTCAAC 2075  
DB 1276 ATTACTGAATCTGCTAGAAATTTTACTGAGTACTTTGGACAGGAAATAAATTTCAAC 1335  
QY 2076 TTCCAGGGTTATAGCCAGGACAAATTTGGAAAGAGTATATGCAAGTAAATATGACCTTAAG 2135  
DB 1336 TTCCAGGGTTATAGCCAGGACAAATTTGGAAAGAGTATATGCAAGTAAATATGACCTTAAG 1395

QY	2136	TATATCTCGATCAGCAAAACCCACAATATGACAGAAAGATTAAAGTGCAGTCTCTTGAA	2195
DB	1396	TATATCTCGATCAGCAAAACCCACAATATGACAGAAAGATTAAAGTGCAGTCTCTTGAA	1455
QY	2196	GGTTTTCGATCTTTTTTGAAGATTCTTACCTGTATGACAGGAATGGAAGAAATCCGAAGA	2255
DB	1456	GGTTTTCGATCTTTTTTGAAGATTCTTACCTGTATGACAGGAATGGAAGAAATCCGAAGA	1515
QY	2256	CAGTTGGGCAACACATTTGAAGTGCATCTGATTTGGGAGGCTGCCATGCTATACAGATG	2315
DB	1516	CAGTTGGGCAACACATTTGAAGTGCATCTGATTTGGGAGGCTGCCATGCTATACAGATG	1575
QY	2316	CAATTTGAAGATAATTTTACTCATGTTTCCAGAGTGTGTGCTGTGTGATGAAGAACTCTTAA	2375
DB	1576	CAATTTGAAGATAATTTTACTCATGTTTCCAGAGTGTGTGCTGTGTGATGAAGAACTCTTAA	1635
QY	2376	CTTGTGGCTTATAAAGAAATGTACAAAGCTGTGATGAGGTGCAGTACAGTTTTCATATCT	2435
DB	1636	CTTGTGGCTTATAAAGAAATGTACAAAGCTGTGATGAGGTGCAGTACAGTTTTCATATCT	1695
QY	2436	AGTAGCAAGACAGTAGTACAAATCGTGTGGACATAGTTTGGAAAAAACAAGTCTCACAGATA	2495
DB	1696	AGTAGCAAGACAGTAGTACAAATCGTGTGGACATAGTTTGGAAAAAACAAGTCTCACAGATA	1755
QY	2496	TCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCTTGTGCTTTCATGTA	2555
DB	1756	TCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCTTGTGCTTTCATGTA	1815
QY	2556	CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGAATGATGAATTTGTGCTTTTTGAGGACTTTT	2615
DB	1816	CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGAATGATGAATTTGTGCTTTTTGAGGACTTTT	1875
QY	2616	CAAGTAGAGGTACTAGTGGAAATATCTTTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT	2675
DB	1876	CAAGTAGAGGTACTAGTGGAAATATCTTTTACGTTGTCTGCTGCTGCTGCTGCTGCTGCT	1935
QY	2676	GCTGAGATGTGGGAAAGAAATGGAAGTCTCTTATTTAGCCAGGTGTTTATTTACCAAGAT	2735
DB	1936	GCTGAGATGTGGGAAAGAAATGGAAGTCTCTTATTTAGCCAGGTGTTTATTTACCAAGAT	1995
QY	2736	GTTAAGTGCAGAGAAATGTATGATTAAGATATCATCATGCTTTCAGATTTGGTGCACT	2795
DB	1996	GTTAAGTGCAGAGAAATGTATGATTAAGATATCATCATGCTTTCAGATTTGGTGCACT	2055
QY	2796	TTAATGGAATCCCAATAGTCTCTGTTTATCTGTTATCTTACAGAGTATGAATCTGCCAGGCT	2855
DB	2056	TTAATGGAATCCCAATAGTCTCTGTTTATCTGTTATCTTACAGAGTATGAATCTGCCAGGCT	2115
QY	2856	TTTAAACAGACCATATCTACAAAAGACCAAGATTTGATTAACCAATATAATACACTAATA	2915
DB	2116	TTTAAACAGACCATATCTACAAAAGACCAAGATTTGATTAACCAATATAATACACTAATA	2175
QY	2916	GAAAGAAATGCTTCAGGCTCATCTATATTTGTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2975
DB	2176	GAAAGAAATGCTTCAGGCTCATCTATATTTGTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2235
QY	2976	AATGTGAACAAAGAGAGGTCAATATGAGAAATCATCTGCTTTGCAATGGAACCC	3035
DB	2236	AATGTGAACAAAGAGAGGTCAATATGAGAAATCATCTGCTTTGCAATGGAACCC	2295
QY	3036	ATGCCACACAGTCCATTTGCCAAAATTTTACTGAGAAATGAAATAATGAAGAACTGGCTTAA	3095
DB	2296	ATGCCACACAGTCCATTTGCCAAAATTTTACTGAGAAATGAAATAATGAAGAACTGGCTTAA	2355
QY	3096	GAGAAATGTCAATAAACAAGTGGCCACATTTTAAAGAACCAAGTGTATCAGGCCATGGAGTT	3155
DB	2356	GAGAAATGTCAATAAACAAGTGGCCACATTTTAAAGAACCAAGTGTATCAGGCCATGGAGTT	2415
QY	3156	TATGAATCAAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAA	3215
DB	2416	TATGAATCAAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAA	2475
QY	3216	ACCCAGCATAGCAAGGCTGNAACATATGCAAGAAAGAGAGAAAAACAAGAAACAAGAT	3275

DB	2476	ACCAGCATAGCAAGGCTGAACATATGCAAGAAAAAGAGAAAAACAAGAAAAACAAGAT	2535
QY	3276	GAAGCATTTGGCGCACACACCTCTCGTAATCTCGCCCTGCTTTTCAGCAAGTGTATTAAAC	3335
DB	2536	GAAGCATTTGGCGCACACACCTCTCGTAATCTCGCCCTGCTTTTCAGCAAGTGTATTAAAC	2595
QY	3336	CTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTGAGCGGGCAATAGAC	3395
DB	2596	CTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTGAGCGGGCAATAGAC	2655
QY	3396	ACAGATTCTAACTTTGTGGACCGGAAGGGATGCTCCAAATGCTTTTCATATTTCTGGCAATG	3455
DB	2656	ACAGATTCTAACTTTGTGGACCGGAAGGGATGCTCCAAATGCTTTTCATATTTCTGGCAATG	2715
QY	3456	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTCGAAGAGAGTAAACATTTGAC	3515
DB	2716	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTCGAAGAGAGTAAACATTTGAC	2775
QY	3516	TTTTATCATTAAGGCTTCAAGATTGGGAAGTTTCAAGCCATGAATATACAAATGCTTTTGAA	3575
DB	2776	TTTTATCATTAAGGCTTCAAGATTGGGAAGTTTCAAGCCATGAATATACAAATGCTTTTGAA	2835
QY	3576	AAACTCAAAAGGAAATTTCCCAAGTTAGAAAGGCGCAGAGCAATGATAAAGTGGATACTTCAG	3635
DB	2836	AAACTCAAAAGGAAATTTCCCAAGTTAGAAAGGCGCAGAGCAATGATAAAGTGGATACTTCAG	2895
QY	3636	ATGTTTGCACAGTGAAGCGATTAAGAGAAAAATCTTTGTTTAAATTTGATAGCAACCATCA	3695
DB	2896	ATGTTTGCACAGTGAAGCGATTAAGAGAAAAATCTTTGTTTAAATTTGATAGCAACCATCA	2955
QY	3696	GGATCGGAATCTATTAAGAATGATGAGATTACTCATGATAAGAAAAAGACAGAAACGAAAA	3755
DB	2956	GGATCGGAATCTATTAAGAATGATGAGATTACTCATGATAAGAAAAAGACAGAAACGAAAA	3015
QY	3756	AGAAAACTGAGCTGTAGGCTACATCGCCAGAGAGATCAATGCTCAGATGCTGCGCTTAA	3815
DB	3016	AGAAAACTGAGCTGTAGGCTACATCGCCAGAGAGATCAATGCTCAGATGCTGCGCTTAA	3075
QY	3816	CAGAAAACTTCAATGAACTCAATAAATCATGATCATGATGATGATGATGATGATGATGATGAT	3875
DB	3076	CAGAAAACTTCAATGAACTCAATAAATCATGATCATGATGATGATGATGATGATGATGATGAT	3135
QY	3876	AAAGAAATTTCCATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3935
DB	3136	AAAGAAATTTCCATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3195
QY	3936	GCTTTGGGTCTTAAACGGGGTCCATCTGTTTCTGAAAGGAGGTGCTGACGTGCAATCCTT	3995
DB	3196	GCTTTGGGTCTTAAACGGGGTCCATCTGTTTCTGAAAGGAGGTGCTGACGTGCAATCCTT	3255
QY	3996	TGCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	4055
DB	3256	TGCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3315
QY	4056	CAGAAATCTCTGCTTAAACCCAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	4115
DB	3316	CAGAAATCTCTGCTTAAACCCAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3375
QY	4116	CTAGAACCCATCTTTCATGGAATCCAGATTTGGCATATGGAATTTATACAGGAAGCTTGCT	4175
DB	3376	CTAGAACCCATCTTTCATGGAATCCAGATTTGGCATATGGAATTTATACAGGAAGCTTGCT	3435
QY	4176	CATGTAATGACGAGAGTGTGCTGCGAGAGTATTTTGAAGCTGTACAGTGTGCTTCAG	4235
DB	3436	CATGTAATGACGAGAGTGTGCTGCGAGAGTATTTTGAAGCTGTACAGTGTGCTTCAG	3495
QY	4236	CAGCGCATCTCATGTTTGAATCTTTTGAATTTGAGTGTGAGAGTATCTTTGCGCTCTTTC	4295
DB	3496	CAGCGCATCTCATGTTTGAATCTTTTGAATTTGAGTGTGAGAGTATCTTTGCGCTCTTTC	3555
QY	4296	AAATCTCTGTGCAATATCTGTGATCCCATTTATCTTTGCAACCTTCAAAAGATATAACAGT	4355



3556 AATCTCTGTGCAATGCTGATCCCATATTCTTTGCAACCTCAAAAGATAAACAGT 3615  
4356 GAGAAATCAGATGCTCTTGCTCAACTTTGACCCCTGGCAGCGTGATACAGACTTTCTG 4415  
3616 GAGAAATCAGATGCTCTTGCTCAACTTTGACCCCTGGCAGCGTGATACAGACTTTCTG 3675  
4416 GCCAAATATCAGGTTATATATATAGACATGCTTAAGGAGAAACCCAAATCTCTATTTTC 4475  
3676 GCCAAATATCAGGTTATATATATAGACATGCTTAAGGAGAAACCCAAATCTCTATTTTC 3735  
4476 TTTAATCAAGGAATGGGAGATTCTACTTTGGAGTTCCATTCATCTCAGATTTTGGCGTT 4535  
3736 TTTAATCAAGGAATGGGAGATTCTACTTTGGAGTTCCATTCATCTCAGATTTTGGCGTT 3795  
4536 GAGTCTTCGATTAATATTAATATAGCATCAAGGAATATGGTTATTTCTCTTTGCCACAACA 4595  
3796 GAGTCTTCGATTAATATTAATATAGCATCAAGGAATATGGTTATTTCTCTTTGCCACAACA 3855  
4596 ATTTATAGATTTGGATTTGAAGTGGCCTGATGAAAGGATCTCTCGATCCCAATGCTG 4655  
3856 ATTTATAGATTTGGATTTGAAGTGGCCTGATGAAAGGATCTCTCGATCCCAATGCTG 3915  
4656 ACCTGGAGCAGCTGGCTTTCACTATCCAGGCAATTTGAAATCTATTTGGAGATGAAGCA 4715  
3916 ACCTGGAGCAGCTGGCTTTCACTATCCAGGCAATTTGAAATCTATTTGGAGATGAAGCA 3975  
4716 AAACCTCTGTTTGGAGCACTTCAAAATAGGAGCAATATGGTCTGAAAGCAATTAATGAG 4775  
3976 AAACCTCTGTTTGGAGCACTTCAAAATAGGAGCAATATGGTCTGAAAGCAATTAATGAG 4035  
4776 TTGAGATTTGCAGAGATTAATCTGCTCAGGTCCTGATACAGAAACATCTGGTTGCT 4835  
4036 TTGAGATTTGCAGAGATTAATCTGCTCAGGTCCTGATACAGAAACATCTGGTTGCT 4095  
4836 CTTCATCAGTTGTTCTTCTTAACATAAATCAGAAATACACCATGCTCTCTGCTATA 4895  
4096 CTTCATCAGTTGTTCTTCTTAACATAAATCAGAAATACACCATGCTCTCTGCTATA 4155  
4896 GATCTGTTTCAATGTTTGGTGGTCTGTTAGCAATCCCATCTCTGTTGGATGAC 4955  
4156 GATCTGTTTCAATGTTTGGTGGTCTGTTAGCAATCCCATCTCTGTTGGATGAC 4215  
4956 CCGTTGATCTGAGCCTTCTTCAGTTAGTTCTTCTATTAACCATCTTATCTCTTCAT 5015  
4216 CCGTTGATCTGAGCCTTCTTCAGTTAGTTCTTCTATTAACCATCTTATCTCTTCAT 4275  
5016 TTGATCACCATGGCACATGCTTCAAGATCTACTTACAGTAGACACAGGCTACCCCTT 5075  
4276 TTGATCACCATGGCACATGCTTCAAGATCTACTTACAGTAGACACAGGCTACCCCTT 4335  
5076 GCTCAGGTTCAAGAGACAGTGAAGGCTCAATCCGATCTCTTCTTTTGCAGAAAT 5135  
4336 GCTCAGGTTCAAGAGACAGTGAAGGCTCAATCCGATCTCTTCTTTTGCAGAAAT 4395  
5136 TCTCAATATACAGTGGCTTCAATGGGTTGATATCTCTGCTGATTTTGGGCTTCA 5195  
4396 TCTCAATATACAGTGGCTTCAATGGGTTGATATCTCTGCTGATTTTGGGCTTCA 4455  
5196 CTGAAGATGCAATCACCCTTATCTCTGCTGCTGCTGATTTTTCATTTACTT 5255  
4456 CTGAAGATGCAATCACCCTTATCTCTGCTGCTGCTGATTTTTCATTTACTT 4515  
5256 GGGGTAACTCCGCTGAGCACTGATACCAATCTGCAAGAGGAGATGACAGTGCATC 5315  
4516 GGGGTAACTCCGCTGAGCACTGATACCAATCTGCAAGAGGAGATGACAGTGCATC 4575  
5316 TGTAGCTATCTATCTTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATACT 5375  
4576 TGTAGCTATCTATCTTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATACT 4635  
5376 GTAAGGCCCTTGTCCAGAGTGGTGTGAGATCTCTGCTTACTAACTGTTTGAAGCA 5435  
4636 GTAAGGCCCTTGTCCAGAGTGGTGTGAGATCTCTGCTTACTAACTGTTTGAAGCA 4695

QY 5436 AAAAACAACGCTGGTCAAGTACCTTAGAAAAAGAAATAGTTTATAGAGCTTCTCATGAC 5495  
Db |||||  
QY 4696 AAAAACAACGCTGGTCAAGTACCTTAGAAAAAGAAATAGTTTATAGAGCTTCTCATGAC 4755  
Db |||||  
QY 5496 TATAGCTCCCTCCCTGAATCAAGCTTCTCATTTCAAGTGGCCCAAGCTCTGAGATGAG 5555  
Db |||||  
QY 4756 TATAGCTCCCTCCCTGAATCAAGCTTCTCATTTCAAGTGGCCCAAGCTCTGAGATGAG 4815  
Db |||||  
QY 5556 CGAAAGCATCTCTGCTCTCTGCTTTTCTGTTGGGGCTATATCTATGTTCTCAGAAACATTTGC 5615  
Db |||||  
QY 4816 CGAAAGCATCTCTGCTCTCTGCTTTTCTGTTGGGGCTATATCTATGTTCTCAGAAACATTTGC 4875  
Db |||||  
QY 5616 TGCCAGGAAATTTGTAACGGGAGAGAGGTTGGAGCTTGGATTTTTCACGCACTTCACTGT 5675  
Db |||||  
QY 4876 TGCCAGGAAATTTGTAACGGGAGAGAGGTTGGAGCTTGGAGCTTGGATTTTTCACGCACTTCACTGT 4935  
Db |||||  
QY 5676 GGAGCGGAGTCTGATTTTCTTAAATCAGAGAAATGCGAGTGGTCTGTTCAAGGT 5735  
Db |||||  
QY 4936 GGAGCGGAGTCTGATTTTCTTAAATCAGAGAAATGCGAGTGGTCTGTTCAAGGT 4995  
Db |||||  
QY 5736 AAAGCCAGAGCTGTGCTCTATCCAGCTCTTACTTGGATGAATATGGAAGAACACAGCCCT 5795  
Db |||||  
QY 4996 AAAGCCAGAGCTGTGCTCTATCCAGCTCTTACTTGGATGAATATGGAAGAACACAGCCCT 5055  
Db |||||  
QY 5796 GGCTGAAGAGGGGCAACCCCTTCAATTTATCTCTGAGCGGTATCGAAAGCTCATTG 5855  
Db |||||  
QY 5056 GGCTGAAGAGGGGCAACCCCTTCAATTTATCTCTGAGCGGTATCGAAAGCTCATTG 5115  
Db |||||  
QY 5856 GTCTGGCAACAACATGCTATATAGAGAGATTTCTAGGCGGAGAGACTAATCAGATG 5915  
Db |||||  
QY 5116 GTCTGGCAACAACATGCTATATAGAGAGATTTCTAGGCGGAGAGACTAATCAGATG 5175  
Db |||||  
QY 5916 TTATTGGATTTCAACTGGCAGTTACTGTGA 5945  
Db |||||  
QY 5176 TTATTGGATTTCAACTGGCAGTTACTGTGA 5205  
Db |||||

RESULT 6  
ADS86880  
ID ADS86880 standard; cDNA; 5205 BP.  
XX  
AC ADS86880;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
DE cDNA encoding human E3alpha ubiquitin ligase, with SNP.  
XX  
KW Human; E3alpha ubiquitin ligase; huE3alpha1;  
KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;  
KW fasting; metabolic acidosis; muscle degeneration; kidney failure;  
KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;  
KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;  
KW inflammatory cachexia; hyperthyroidism; denervation atrophy;  
KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;  
KW gene therapy; ss; gene; SNP; single nucleotide polymorphism.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 1. 5205  
FT CDS /\*tag= a  
FT /product= "E3alpha ubiquitin ligase"  
XX  
XX US2004185037-A1.  
XX  
PD 23-SEP-2004.  
XX  
XX 15-JAN-2004; 2004US-00758672.  
XX  
XX 08-MAR-2000; 2000US-0187911P.  
PR 28-NOV-2000; 2000US-00724126.  
XX

PA (HANH/) HAN H.  
XX (KWAK/) KWAK K.  
PI Han H, Kwak K;  
XX WPI; 2004-707854/69.  
XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule  
PI useful for treating and/or preventing renal cachexia or inflammatory  
XX cachexia.  
PS Example 9; SEQ ID NO 18; 115pp; English.  
XX  
CC The present invention relates to new orthologue of human E3alpha  
CC ubiquitin ligase, huE3alpha1 and huE3alpha2. Most intracellular proteins  
CC are degraded through the ubiquitin-proteasome pathway. Proteins are  
CC marked for proteasomal degradation by conjugation of ubiquitin to the  
CC protein. Conjugation of the ubiquitin molecule involves the activation by  
CC E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,  
CC and then interacts with a specific E3 ligase family member. E3 ligase  
CC binds to proteins targeted for degradation and catalyses the transfer of  
CC ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase  
CC determines the specificity of the system. The E3alpha family is the main  
CC family of intracellular ligases and is involved in the N-end rule pathway  
CC of protein degradation. E3alpha enzyme binds directly to the primary  
CC destabilising N-terminal amino acid and catalyses ubiquitin conjugation  
CC thereby targeting the protein for degradation. The human E3alpha gene is  
CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-  
CC proteasome pathway has been determined to be a major cause of rapid  
CC muscle wasting including, fasting, metabolic acidosis, muscle  
CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,  
CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen  
CC balance, burns, Cushing's syndrome, inflammatory cachexia,  
CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-  
CC protein malnutrition. E3alpha plays a role in the overall increase in  
CC ubiquitination that is associated with and may mediate muscle atrophy in  
CC catabolic and other disease states. Treatment may be administered by gene  
CC therapy, cell therapy and antisense therapy methods. The present sequence  
CC is an assembled sequence of human E3alpha ubiquitin ligase. The patent  
CC describes this sequence as having a SNP at nucleotide 5397 of the  
CC sequence causing a change of cytosine to a thymidine at position 4702 of  
CC human E3alpha1. However, this sequence is shown in the sequence listing  
CC to have an adenine at position 4702, so no features could be done.  
XX  
SQ Sequence 5205 BP; 1553 A; 1017 C; 1171 G; 1450 T; 0 U; 14 Other;  
Query Match 81.4%; Score 5136; DB 13; Length 5205;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;  
QY 696 ATGCGGACGAGGAGGCTGGAGGTAATCTGAGAGGATGGAATCAGCGCGAGGTTACCCGAG 755  
DB 1 ATGCGGACGAGGAGGCTGGAGGTAATCTGAGAGGATGGAATCAGCGCGAGGTTACCCGAG 60  
QY 756 ACCCTCAGCGCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTCTTG 815  
DB 61 ACCCTCAGCGCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTCTTG 120  
QY 816 CATCATTTGGGCAATTTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 875  
DB 121 CATCATTTGGGCAATTTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 180  
QY 876 AAGCAGGAGGAAAGTGTAACAATGTCATATTTCACTCCACTCGGAATGGTACTTATTTGGA 935  
DB 181 AAGCAGGAGGAAAGTGTAACAATGTCATATTTCACTCCACTCGGAATGGTACTTATTTGGA 240  
QY 936 GAAGATCCAGATATTTGCTTAGAGAAATTTAGAGCAAGTGGAGCAATTTTCAGCTTTGTGG 995  
DB 241 GAAGATCCAGATATTTGCTTAGAGAAATTTAGAGCAAGTGGAGCAATTTTCAGCTTTGTGG 300  
QY 996 AGGGTTTTCAAAAGTGGAGAGCAACCTATTCTTTCAGGAGATTTGCAATTTGATCCAAACA 1055  
DB 301 AGGGTTTTCAAAAGTGGAGAGCAACCTATTCTTTCAGGAGATTTGCAATTTGATCCAAACA 360

QY 1056 TGTGTACTCTGTATGGACTGCTTCCAGGACAGTGTTCATATAAATCATCGTTTCAAGATG 1115  
DB 361 TGTGTACTCTGTATGGACTGCTTCCAGGACAGTGTTCATATAAATCATCGTTTCAAGATG 420  
QY 1116 CATACTTCTACTGAGGAGGAGGTTCTGTGACTGTGTGGAGACACAGAGGCGATGGAATCGGC 1175  
DB 421 CATACTTCTACTGAGGAGGAGGTTCTGTGACTGTGTGGAGACACAGAGGCGATGGAATCGGC 480  
QY 1176 CCTTTTGTGTAAATCATGAACTGTGAAAGAGCGGTACTATATAAAGAGAAATTCAGCTGT 1235  
DB 481 CCTTTTGTGTAAATCATGAACTGTGAAAGAGCGGTACTATATAAAGAGAAATTCAGCTGT 540  
QY 1236 CCGTTGTAATGAGAGGTAAATTTGTCCAAAGCCAGGAAATATTTCTCTCAGTGATATAATAT 1295  
DB 541 CCGTTGTAATGAGAGGTAAATTTGTCCAAAGCCAGGAAATATTTCTCTCAGTGATATAATAT 600  
QY 1296 GTCTGTAGAAATGACTATATGCGAAGAGGAAAGAACTGCTCTCTGAACTCCAGATAAGG 1355  
DB 601 GTCTGTAGAAATGACTATATGCGAAGAGGAAAGAACTGCTCTCTGAACTCCAGATAAGG 660  
QY 1356 GAGAAAAATGAAAGATACTATTTGTCTTTCAATGATGAACACCATTCATATGACAC 1415  
DB 661 KNR -----YTCVNDHHSYDH 675  
QY 1416 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTGCATACC 1475  
DB 676 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTGCATACC 735  
QY 1476 ACTGCCATTGCAAAAGAGGCTGTCGGGCTGTTTAAAGCGGAGCTTATGCTGCTTGCAG 1535  
DB 736 ACTGCCATTGCAAAAGAGGCTGTCGGGCTGTTTAAAGCGGAGCTTATGCTGCTTGCAG 795  
QY 1536 GAAGCAAGGAGGATATAAGAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 1595  
DB 796 GAAGCAAGGAGGATATAAGAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 855  
QY 1596 GAAGTATTACACTCAGAGATTTAGGCTCAGCAAAATTTGCTTGGCTTCTGTTCTTCTGG 1655  
DB 856 GAAGTATTACACTCAGAGATTTAGGCTCAGCAAAATTTGCTTGGCTTCTGTTCTTCTGG 915  
QY 1656 ATGAAACAAAATTTATGAGCTATTTCAAGTGACTTTTAGGCGAGATCTTTTGGCAAGCATGCC 1715  
DB 916 ATGAAACAAAATTTATGAGCTATTTCAAGTGACTTTTAGGCGAGATCTTTTGGCAAGCATGCC 975  
QY 1716 AGAGAGAACCTGACTCGGAGAAATCCCTGTCTCATAGAGAGGTTAATGCTTTGGATGCA 1775  
DB 976 AGAGAGAACCTGACTCGGAGAAATCCCTGTCTCATAGAGAGGTTAATGCTTTGGATGCA 1035  
QY 1776 AAGCTTTTATAAAGGTGCCGTAAGATCTTTCATGAATTTGATCTTCAGCAGTTTTTTTATG 1835  
DB 1036 AAGCTTTTATAAAGGTGCCGTAAGATCTTTCATGAATTTGATCTTCAGCAGTTTTTTTATG 1095  
QY 1836 GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTTAATAAACAATGCGAG 1895  
DB 1096 GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTTAATAAACAATGCGAG 1155  
QY 1896 AAAGAAATATATCAGTGATGATGACAGAAAGTATCTCTATAACTGCACTTTCAGTTTCAG 1955  
DB 1156 AAAGAAATATATCAGTGATGATGACAGAAAGTATCTCTATAACTGCACTTTCAGTTTCAG 1215  
QY 1956 ATGTTTACTGTTCTTACTCTGCGCTCGACATCTTATTTGAAGAGCAGAAATGTTTATCTCTGTC 2015  
DB 1216 ATGTTTACTGTTCTTACTCTGCGCTCGACATCTTATTTGAAGAGCAGAAATGTTTATCTCTGTC 1275  
QY 2016 ATTAAGTAACTCTGCTAGAGGTTTCTGCTGAGTACTTGGACAGGAAACAATAAATTTCAAC 2075  
DB 1276 ATTAAGTAACTCTGCTAGAGGTTTCTGCTGAGTACTTGGACAGGAAACAATAAATTTCAAC 1335  
QY 2076 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTAAATGTCACCTAAG 2135  
DB 1336 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTAAATGTCACCTAAG 1395

QY	2136	TATATCTGATCAGCAGACCCCAATATGACAGAAAGATTAAAGATGAGTTCCTTGAA	2195
DB	1396	TATATCTGATCAGCAAAACCCCAATATGACAGAAAGATTAAAGATGAGTTCCTTGAA	1455
QY	2196	GGTTTTCGATCTTTTTCGAAGATCTTACCTGATGACAGGGAATGGAAGAAATCCGAAGA	2255
DB	1456	GGTTTTCGATCTTTTTCGAAGATCTTACCTGATGACAGGGAATGGAAGAAATCCGAAGA	1515
QY	2256	CAGGTGCGGCAACACATTTGAAGTGGATCCTGATGCGGAGGCTGCCATTTGCTATACAGATG	2315
DB	1516	CAGGTGCGGCAACACATTTGAAGTGGATCCTGATGCGGAGGCTGCCATTTGCTATACAGATG	1575
QY	2316	CAATTTGAAGATATTTTACTCATGTTCCAGAGTGGTGTGCTTGATGAGAACTCTTTA	2375
DB	1576	CAATTTGAAGATATTTTACTCATGTTCCAGAGTGGTGTGCTTGATGAGAACTCTTTA	1635
QY	2376	CTTGTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGTGCAGTACAGATTTTCATATCT	2435
DB	1636	CTTGTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGTGCAGTACAGATTTTCATATCT	1695
QY	2436	AGTAGCAAGACAGTAGTACAAATCGTGTGACACATAGTTTGGAAACAAAGTCTCTACAGATG	2495
DB	1696	AGTAGCAAGACAGTAGTACAAATCGTGTGACACATAGTTTGGAAACAAAGTCTCTACAGATG	1755
QY	2496	TCTGAGGATCTGTGAAGCATACATCTGCCACTCTCTAGACCCCTGCTGCTTCATGTA	2555
DB	1756	TCTGAGGATCTGTGAAGCATACATCTGCCACTCTCTAGACCCCTGCTGCTTCATGTA	1815
QY	2556	CGTTTAAAGCAGCTGGTGTGTTTCAAGACTGCAATGATTTGTCTTTTGGAGACTTT	2615
DB	1816	CGTTTAAAGCAGCTGGTGTGTTTCAAGACTGCAATGATTTGTCTTTTGGAGACTTT	1875
QY	2616	CAAGTAGAGGTACTAGTGGATATATCCTTTACGTTGCTGGTGTGGTTCGCCAGGTGTT	2675
DB	1876	CAAGTAGAGGTACTAGTGGATATATCCTTTACGTTGCTGGTGTGGTTCGCCAGGTGTT	1935
QY	2676	GCTGAGATGTGGCGAAGAAATGGAAGTGTCTTTATAGCCAGGTGTTTATTAACAAGAT	2735
DB	1936	GCTGAGATGTGGCGAAGAAATGGAAGTGTCTTTATAGCCAGGTGTTTATTAACAAGAT	1995
QY	2736	GTTAAGTGCAGAGAAATGATGATAAAGATATCATGCTTCAGATTTGGTGCATCT	2795
DB	1996	GTTAAGTGCAGAGAAATGATGATAAAGATATCATGCTTCAGATTTGGTGCATCT	2055
QY	2796	TTAATGGATCCCAATAGTCTTGTGTTACTGATGATGATGAAATGAACTTGGCCAGGCT	2855
DB	2056	TTAATGGATCCCAATAGTCTTGTGTTACTGATGATGATGAAATGAACTTGGCCAGGCT	2115
QY	2856	TTTAAACAAGACATATCTACAAAGACAGAGTTGATTAACAATATAACTAATA	2915
DB	2116	TTTAAACAAGACATATCTACAAAGACAGAGTTGATTAACAATATAACTAATA	2175
QY	2916	GAGAAATGCTTCAGGCTCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2975
DB	2176	GAGAAATGCTTCAGGCTCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2235
QY	2976	AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCC	3035
DB	2236	AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCC	2295
QY	3036	ATGCCACACAGTGCCATTTGCAAAATTTTACCTGAGAAATGAAATAATGAACTGGCTTA	3095
DB	2296	ATGCCACACAGTGCCATTTGCAAAATTTTACCTGAGAAATGAAATAATGAACTGGCTTA	2355
QY	3096	GAGATGTCTAATAAAGAGTGGCCACATTTAAGAAACAGGTGTATCAGGCCATGGAGTT	3155
DB	2356	GAGATGTCTAATAAAGAGTGGCCACATTTAAGAAACAGGTGTATCAGGCCATGGAGTT	2415
QY	3156	TATGAATCAAAAGAGTGAATCACTGAAAGACTTCAATATGTAATTTTATCTTCTCCAAA	3215
DB	2416	TATGAATCAAAAGAGTGAATCACTGAAAGACTTCAATATGTAATTTTATCTTCTCCAAA	2475
QY	3216	ACCCAGCATAGCAGGCTGAAACATATGCAAGAAAGAGGAGAAACCAAGAT	3275

DB	2476	ACCCAGCATAGCAGGCTGAAACATATGCAAGAAAGAGGAGAAACCAAGAAACCAAGAT	2535
QY	3276	GAGCATTTCGCGCCACACACCTCTCTGAAATTCCTGCTCTGCTTTCAGCAAAAGTGAATTAAC	3335
DB	2536	GAGCATTTCGCGCCACACACCTCTCTGAAATTCCTGCTCTGCTTTCAGCAAAAGTGAATTAAC	2595
QY	3336	CTTCTCAAATCTGTGATATCATGATGATGATTCCTCAGGACCGTATTTGAGCGGCAATAGAC	3395
DB	2596	CTTCTCAAATCTGTGATATCATGATGATGATTCCTCAGGACCGTATTTGAGCGGCAATAGAC	2655
QY	3396	ACAGATTTCTAACTTGTGGACCGAAGGATGCTCCAAATGGCTTTTCATATTCCTGCGATG	3455
DB	2656	ACAGATTTCTAACTTGTGGACCGAAGGATGCTCCAAATGGCTTTTCATATTCCTGCGATG	2715
QY	3456	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTCTGAAAGAAAGATTAACATTTGAC	3515
DB	2716	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTCTGAAAGAAAGATTAACATTTGAC	2775
QY	3516	TTTTATCATAAAGCTTCAAGATTTGGAGATTCAGCATGAATATACAAATGCTTTTGGAA	3575
DB	2776	TTTTATCATAAAGCTTCAAGATTTGGAGATTCAGCATGAATATACAAATGCTTTTGGAA	2835
QY	3576	AAACTCAAGGAATTTCCCGAGTTAGAGGCGCAGAGGACATGATTAACCTGATCTTCAG	3635
DB	2836	AAACTCAAGGAATTTCCCGAGTTAGAGGCGCAGAGGACATGATTAACCTGATCTTCAG	2895
QY	3636	ATGTTTGACACAGTGAAGCGATTAAGAGAAATCTCTGTTTAAATGTTAGCAACCAATCA	3695
DB	2896	ATGTTTGACACAGTGAAGCGATTAAGAGAAATCTCTGTTTAAATGTTAGCAACCAATCA	2955
QY	3696	GGATCGGAATCTATTAAGAATGATGAGATTAATCATGATTAAGAAAGAGCAAGCAAAA	3755
DB	2956	GGATCGGAATCTATTAAGAATGATGAGATTAATCATGATTAAGAAAGAGCAAGCAAAA	3015
QY	3756	AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGATCATGGCTCAGATGCTGCTTGA	3815
DB	3016	AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGATCATGGCTCAGATGCTGCTTGA	3075
QY	3816	CAGAAACTTCTATTAAGAACTCATAACTCATGATGATGATGATGATGATGATGATGATG	3875
DB	3076	CAGAAACTTCTATTAAGAACTCATAACTCATGATGATGATGATGATGATGATGATGATG	3135
QY	3876	AAAGAGATTCATTTATGAGGAGAGAGACCCAGCAGCTCAGTACTCTAGAAAT	3935
DB	3136	AAAGAGATTCATTTATGAGGAGAGAGACCCAGCAGCTCAGTACTCTAGAAAT	3195
QY	3936	GCTTTGGGTCTTAAACCGGGTCCATCTGTTACTGAAAGGAGGCTGCTGACGTCATCTT	3995
DB	3196	GCTTTGGGTCTTAAACCGGGTCCATCTGTTACTGAAAGGAGGCTGCTGACGTCATCTT	3255
QY	3996	TGCCAAGAGAAACAGGAGGTGAAATAGAAATATGCAATGATGATGATGATGATGATGATG	4055
DB	3256	TGCCAAGAGAAACAGGAGGTGAAATAGAAATATGCAATGATGATGATGATGATGATGATG	3315
QY	4056	CAGAAATCTACTGCTTAAACCCAGCAGAGGAGGAAACCCATAGAACTCTCAGGAGAGCC	4115
DB	3316	CAGAAATCTACTGCTTAAACCCAGCAGAGGAGGAAACCCATAGAACTCTCAGGAGAGCC	3375
QY	4116	CTAGACCCCATTTTTCATGAGATTCAGATCTTGGCATATGGAACCTTATACAGGAGCTGCT	4175
DB	3376	CTAGACCCCATTTTTCATGAGATTCAGATCTTGGCATATGGAACCTTATACAGGAGCTGCT	3435
QY	4176	CATGTAATGACGACGATGCTGCGAGAGTATTTTGAAGCTGTACAGCTCAGCTCTCAG	4235
DB	3436	CATGTAATGACGACGATGCTGCGAGAGTATTTTGAAGCTGTACAGCTCAGCTCTCAG	3495
QY	4236	CAGCGCATTCATGTTGACCTTTTTCATGAAAGTGGAGATATCTTTGCTCTCTTTCG	4295
DB	3496	CAGCGCATTCATGTTGACCTTTTTCATGAAAGTGGAGATATCTTTGCTCTCTTTCG	3555
QY	4296	AAATCTCTGTGCAATATCTGTGATCCCATATTTCTTTGCAACCTCAAAAGATTAACAGT	4355

Db 3556 AATCTCTGTGCAATPACTGTGATCCCATTTATCTCTTTCACACCTCAAAAGATAAACAGT 3615  
Qy 4356 GAGATGACAGATGCTCTTGTCTCACTTTTGTACCTTGGCACCGGTGATACAGACTGTCTG 4415  
Db 3616 GAGAATGACAGATGCTCTTGTCTCACTTTTGTACCTTGGCACCGGTGATACAGACTGTCTG 3675  
Qy 4416 GCCAGATATCAGGGTTAATAATATAAGACATGCTAAAGGAGAGAAAACCCCAATTCCTATTTTC 4475  
Db 3676 GCCAGATATCAGGGTTAATAATATAAGACATGCTAAAGGAGAGAAAACCCCAATTCCTATTTTC 3735  
Qy 4476 TTTAATCAAGGAATGGAGATCTTACTTTTGAGATTCATCCATCCCTGAGTTTGGCGTT 4535  
Db 3736 TTTAATCAAGGAATGGAGATCTTACTTTTGAGATTCATCCATCCCTGAGTTTGGCGTT 3795  
Qy 4536 GAGTCTTCGATTAATATTTCAATAGCATCAAGGAATGGTTATCTCTTTCGCCACACAA 4595  
Db 3796 GAGTCTTCGATTAATATTTCAATAGCATCAAGGAATGGTTATCTCTTTCGCCACACAA 3855  
Qy 4596 ATTTATAGAAATGGATTTGAAAGTGCACCTCATGAAAGGGATCCTCGAGTCCCATGCTG 4655  
Db 3856 ATTTATAGAAATGGATTTGAAAGTGCACCTCATGAAAGGGATCCTCGAGTCCCATGCTG 3915  
Qy 4656 ACCTGGAGCACTCGGCTTTCACTATCCAGGCAATGAAATCTATTTGGGAGATGAAGGA 4715  
Db 3916 ACCTGGAGCACTCGGCTTTCACTATCCAGGCAATGAAATCTATTTGGGAGATGAAGGA 3975  
Qy 4716 AAACCTCTGTTGGAGCACTTCAAAATAGGAGCATATGCTCTGAAGCATTAATGCAG 4775  
Db 3976 AAACCTCTGTTGGAGCACTTCAAAATAGGAGCATATGCTCTGAAGCATTAATGCAG 4035  
Qy 4776 TTTGCAAGTTCACAGAGATTAACCTCTCAGGTCCTGATACAGAAACATCTGTTGCT 4835  
Db 4036 TTTGCAAGTTCACAGAGATTAACCTCTCAGGTCCTGATACAGAAACATCTGTTGCT 4095  
Qy 4836 CTCTCATCAGTTGTTCTTCTTAACATAAAATCAGAAATACACCATGCTTCTGTCTATA 4895  
Db 4096 CTCTCATCAGTTGTTCTTCTTAACATAAAATCAGAAATACACCATGCTTCTGTCTATA 4155  
Qy 4896 GATCTGTTTCATGTTTGGTGGTGTCTGTGATGATTTCCCATCTTGTATGGATGAC 4955  
Db 4156 GATCTGTTTCATGTTTGGTGGTGTCTGTGATGATTTCCCATCTTGTATGGATGAC 4215  
Qy 4956 CCGTGTGATCTGCAGCCTTCTCAGTATGTTCTTCTTAACACCATCTTATCTCTTCCAT 5015  
Db 4216 CCGTGTGATCTGCAGCCTTCTCAGTATGTTCTTCTTAACACCATCTTATCTCTTCCAT 4275  
Qy 5016 TTGATCACCATGGCACATGCTTCAGATACCTACTTACAGTAGACACAGGCTACCCCTT 5075  
Db 4276 TTGATCACCATGGCACATGCTTCAGATACCTACTTACAGTAGACACAGGCTACCCCTT 4335  
Qy 5076 GCTCAGGTTCAAGAGACAGTGAAGAGCTCATTCGGCATCTTCTTCTTTCGAGAAAT 5135  
Db 4336 GCTCAGGTTCAAGAGACAGTGAAGAGCTCATTCGGCATCTTCTTCTTTCGAGAAAT 4395  
Qy 5136 TCTCAATATACAAGTGGCTCCATTTGGGTGATATCTCGCTGGTGTATTTTGGGTCTCA 5195  
Db 4396 TCTCAATATACAAGTGGCTCCATTTGGGTGATATCTCGCTGGTGTATTTTGGGTCTCA 4455  
Qy 5196 CTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGCTGATTTGTTTTCACATTTACTT 5255  
Db 4456 CTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGCTGATTTGTTTTCACATTTACTT 4515  
Qy 5256 GGGTAACTCCGCTGAGAACTGATACCAATTCGCAAGAGAGATACAGTGCATC 5315  
Db 4516 GGGTAACTCCGCTGAGAACTGATACCAATTCGCAAGAGAGATACAGTGCATC 4575  
Qy 5316 TGTAGCTATCTATCTTTACCTACAATTTGTTCTTCTTCCAGGAATATGGGATACT 5375  
Db 4576 TGTAGCTATCTATCTTTACCTACAATTTGTTCTTCTTCCAGGAATATGGGATACT 4635  
Qy 5376 GTAAGGCCCTTCTCCAGAGTGGTGTGCAGATCTGCTTACTAAACCTGTTTGAAGCAA 5435  
Db 4636 GTAAGGCCCTTCTCCAGAGTGGTGTGCAGATCTGCTTACTAAACCTGTTTGAAGCAA 4695

Qy 5436 AAAAAACCCGTGGTACGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGAC 5495  
Db 4696 AAAAAACCCGTGGTACGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGAC 4755  
Qy 5496 TATAGTCCCTCTCGAATCAAGCTTCTCAATTCAGGTGCCACCGTCTGCAAGATGAG 5555  
Db 4756 TATAGTCCCTCTCGAATCAAGCTTCTCAATTCAGGTGCCACCGTCTGCAAGATGAG 4815  
Qy 5556 CGAAGCATCTGTCTCTGCTTCTGTTGGGCTATATCTATCTCTCAGAAATTTGC 5615  
Db 4816 CGAAGCATCTGTCTCTGCTTCTGTTGGGCTATATCTATCTCTCAGAAATTTGC 4875  
Qy 5616 TGCCAGCAAAATTTGAACGGGAAAGAGTTGGAGCTTTGCAATTTTTCAGCACTTCACTGT 5675  
Db 4876 TGCCAGCAAAATTTGAACGGGAAAGAGTTGGAGCTTTGCAATTTTTCAGCACTTCACTGT 4935  
Qy 5676 GGAGCCGAGTCTGCAATTTTCTAAAAATCAGAAATGCGAGTGGTCTCGTTGAAGGT 5735  
Db 4936 GGAGCCGAGTCTGCAATTTTCTAAAAATCAGAAATGCGAGTGGTCTCGTTGAAGGT 4995  
Qy 5736 AAAGCCAGAGCTGTGCTATCCAGCTCTTACTTGGATGAAATATGGAGAAACAGACCT 5795  
Db 4996 AAAGCCAGAGCTGTGCTATCCAGCTCTTACTTGGATGAAATATGGAGAAACAGACCT 5055  
Qy 5796 GGCCTGAAGAGGGGCAACCCCTTCAATTTATCTCTGAGCGGTATCGGAAGCTCCATTTG 5855  
Db 5056 GGCCTGAAGAGGGGCAACCCCTTCAATTTATCTCTGAGCGGTATCGGAAGCTCCATTTG 5115  
Qy 5856 GTCTGGCAACAACACTCATTTATAGAAGATCTCAGGAGCCAGAGACTAATCAGATG 5915  
Db 5116 GTCTGGCAACAACACTCATTTATAGAAGATCTCAGGAGCCAGAGACTAATCAGATG 5175  
Qy 5916 TTATTTGGATTTCACTGGCAGTTACTGTGA 5945  
Db 5176 TTATTTGGATTTCACTGGCAGTTACTGTGA 5205

## RESULT 7

AAV99308  
ID AAV99308 standard; cdna; 6395 BP.

XX AAV99308;

AC AAV99308;

DT 25-MAR-1999 (first entry)

XX cDNA encoding mouse a ubiquitin-protein ligase, Ubr1.

DE Ubiquitin-protein ligase; Ubr1; mouse; ubiquitinylation; degradation;

XX N-end rule pathway; stress-related muscle wasting; inhibitor; screen; ss.

OS Mus sp.

XX Key Location/Qualifiers

XX CDS 115..5388

XX FT /\*tag= a

XX FT /product= "Ubr1"

XX US5861312-A.

XX 19-JAN-1999.

XX 02-DEC-1997; 97US-00982956.

XX 02-DEC-1997; 97US-00982956.

XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.

XX Kwon YT, Varshavsky A;

XX WPI; 1999-130395/11.

XX P-PSDB; AAW84351.

PT Mouse and human Ubri1 cDNA - useful for producing recombinant Ubri1  
XX polypeptides.

PS Claim 1; Col 15-28; 18pp; English.

XX The present sequence encodes a ubiquitin-protein ligase called Ubri1. The  
CC Ubri1 enzymes are involved in protein ubiquitinylation and ultimate  
CC degradation through the N-end rule pathway and have been linked to stress  
CC -related muscle wasting. Recombinant Ubri1 polypeptides can be used to  
CC screen for inhibitors of muscle wasting when this is associated with the  
CC N-end rule pathway

XX SQ Sequence 6395 BP; 1802 A; 1376 C; 1519 G; 1698 T; 0 U; 0 Other;

Query Match 70.3%; Score 4435.4; DB 2; Length 6395;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;

QY	590	TTCCAGGGCGCGTAAAGTGTCTCCCTGCTCTC-TCCGACGGGCCACAGGTTTCCGCT	648
DB	8	TTCCAGGGCGCGTAAAGTGTCTCCCTGCTCTGCGGTGCGGCCGCCACAGGTTTCCGCT	67
QY	649	TGCTCTCTGCGCGGGGTGGCAACTGCGAGCGTCAAGTTTCCCTCAAGATGGCGACGAGG	708
DB	68	AGCTGCGCGCGGGGTGGCAACTGCGGGCGTGTCTTCCCTTAAGATGGCGACGAG	127
QY	709	AGGCTGGAGGTTACTGAGAGATGGAATCAAGCGCGAGTTACCCAGACCCCTCAGCGTC	768
DB	128	AGATGGACGGCGCGGAGGATGGACGTGAGCGCGGAGCCTCCCTGCGCGCGACGCGC	187
QY	769	TGGCATCTGGTGGGATCAGCAAGTTGATTTTATCTACTGCTTCTTGATCATTTGGCAC	828
DB	188	CGGCATCTGGTGGGATCAGCAAGTTGATTTTATCTACTGCTTCTTGATCATTTGGCAC	247
QY	829	AATTCGTGCGCAAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGAGGAAA	888
DB	248	AATTAAGTCCAGAAATTTATTTGCTGAGATGGAACCCAGATTTGGAAAGCAGAGAGA	307
QY	889	GTGTACAAATGTCATATTTCACTCCACTGGAATGGTACTTATTTGGAGAAGATCCAGATA	948
DB	308	GTGTACAGATGTCATATTTCACTCCCTTTGGAGTGTACTTATTTGGAGAGATCCGATA	367
QY	949	TTTGCTTAGAGAAATTTGAAGCAGTGGAGCAATTTGAGCTTTGCGGAGGGTTTCAAAA	1008
DB	368	TTTGCTTAGAGAAATTTAAACAACAGTGGAGCGTCCAGTTGTGTGGAGGGTTTCAAAA	427
QY	1009	GTGGAGAGCAACCTATTCTTCTGACGGATTTGCAATTTGATCCCAATGTGTACTCTGTA	1068
DB	428	GTGGAGAAACACATATTCTCTGAGGGATTTGCAATTTGATCCCAATGTGTACTCTGTA	487
QY	1069	TGGAATGCTCCAGAGCAGTGTTCATAAAATCATCGTTTACAGATGCATCTTCTACTG	1128
DB	488	TGGAATGCTCCAGAGTGTTCATAAAATCATCGTTTACAGATGCATCTTCTACTG	547
QY	1129	GAGGAGGGTTCTGTGACTGTGGAGACACAGAGGCATGGAATCTGGCCCTTTTGTGTAA	1188
DB	548	GAGGAGGGTTCTGTGACTGTGGAGACACAGAGGCATGGAATCTGGCCCTTTTGTGTGG	607
QY	1189	ATCATGAACCTGGAGAGCGTACTATTAAGAGAAATTCACGCTGTCGGTGAATGAAG	1248
DB	608	ATCATGAACCTGGAGAGCGTACTATTAAGAGAAATTCACGCTGTCGGTGAATGAAG	667
QY	1249	AGGTAAATGTCCAGCCAGGAAATATTTCTTTCAGTGATAAAATATGTCGTAGAAATGA	1308
DB	668	AGGTAAATGTCCAGCCAGGAAATATTTCTTTCAGTGATAAAATATGTCGTAGAAATGA	727
QY	1309	CTATATGGAGAGAGAAAGAACTGCTCTCTGAACTCCAGATAGGAGAGAAATGAAG	1368
DB	728	CTATATGGAGAGAGAAAGAAATGCTCTCTGAACTCCAGATAGGAGAGAAATGAAG	787
QY	1369	GATACATATGTCCTTTTCAATGATGAACCACTTATATGACCACTCATATACAGCC	1428
DB	788	GATACATATGTCCTTTTCAACGATGAGCACCATTCTGATATGATCATGTGATCTACAGTC	847

QY	1429	TACAAAGAGCTCTTGACTGTGAGCTGCGCAGAGGCCAGCTGTGCATACCACTGCGCATTGACA	1488
DB	848	TGCAGAGAGCTCTAGATTGCGAGCTTGCGAGAGGCACAGCTGCACAGCTGCGCATCGACA	907
QY	1489	AAGAGGGTCGTGCGGCTGTAAAGCGGAGCTTATGCTGCTTGCAGGAAGCAAAAGGAG	1548
DB	908	ANGAGGGTCGCGGGCTGTCAAGCAGAGTGTGTATGCCACTTGCAGGAAGCAAAAGGAG	967
QY	1549	ATATAAAGAGTCATTTCAGAAAATGCTCTCAACATCCCACTTTCATGTAGAAATGATTACAT	1608
DB	968	ATATAAAGAGTCATTTCAGAAAATGCTCTCAACATCCCACTTTCATGTAGAAATGATTACAT	1027
QY	1609	CAGAGATTATGGCTCATCAGAAAATTTGCTTTGCGTCTTGGTTCCTGGATGAACAAAATTA	1668
DB	1028	CCGTGGTTATGGCTCACCAGAAAATTCGCTCTGCGCTTGGCTCTCTGGATGAACAAAATTA	1087
QY	1669	TGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCACAGCATGCTTTCAGGAAGAACCTG	1728
DB	1088	TGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCACAGCATGCTTTCAGGAAGAACCTG	1147
QY	1729	ACTCGGAAATCCCTGCTCATAGCAGGTTAAATGCTTTGGGATGCAAAAGCTTTATAAAG	1788
DB	1148	GCTCTGAAATCCCTGCTCATAGCAGCATTAATGCTTTGGGATGCAAAATCTTTATAAAG	1207
QY	1789	GTGCCGTGAAGATCCTTCATGAAATTTGATCTTCAGCAGTCTTTTATGAGAGATGGAATACA	1848
DB	1208	GTGCCGTGAAGATCCTTCATGAAATTTGATCTTTAGTAGTCTTTTATGAGAGATGGAATACA	1267
QY	1849	AAAAACTCTTGTATGGAATTTGAGTAGTATTAATAACAATGCGAGAAAGATATATCA	1908
DB	1268	AAAAACTCTTGTATGGAATTTGAGTAGTATTAATAACAATGCGAGAAAGATATATCA	1327
QY	1909	GTGATGATCATGACAGAGATATCTATAAATGCACTTTTCAGTTTCAGATGTTTATCTGTTTC	1968
DB	1328	CGCAGCACCAAGAGAGATCTCCATAACCGCCCTGTCGTCGAGATGCTCACCGTCC	1387
QY	1969	CTACTCTGGCTCGACATCTTATGAGAGCAGAAATGTTATCTCTGTCAATTTACTGAAATCTC	2028
DB	1388	CGACTTGGCCCGGCATCTTATGAGAGCAGAAATGTTATTTCTGTCTACTTACTGAAAGCGC	1447
QY	2029	TGCTAGAGATTTTACCTGAGTACTTGGACAGGACAAATTAATTCAACTTCCAGGGTTATA	2088
DB	1448	TGCTAGAGATTTTACCTGAGTACTTGGACAGGACAAATTAATTCAACTTCCAGGGTTATA	1507
QY	2089	GCCAGGACAAATTTGGGAAGAGTATATGAGTAATATGTGACTTAAAGTATATCTCTGATCA	2148
DB	1508	GCCAGGACAAATTTGGGAAGAGTATATGAGTAATATGTGACTTAAAGTATATCTCTGATTA	1567
QY	2149	GCAAAACCCCAATATGGACAGAAATTAAGAAATGCAATGCAATGCAATGCAATGCAATGCA	2208
DB	1568	GCAAGCCTGTCTATGGACAGAAATTAAGAAATGCAATGCAATGCAATGCAATGCAATGCA	1627
QY	2209	TTTTGAGATTTCTTACCTGTATGAGGAAATGGAAGAAATCCGAGACAGAGTTGGGCAAC	2268
DB	1628	TTCTGAGATTTCTTACCTGTATGAGGAAATGGAAGAAATCCGAGACAGAGTTGGGCAAC	1687
QY	2269	ACATTGAAGTGGATCTGATTTGGGAGGCTGCAATTTGCTATACAGATGCAATTTGAAGATA	2328
DB	1688	ACATTGAAGTGGATCTGATTTGGGAGGCTGCAATTTGCTATACAGATGCAATTTGAAGATA	1747
QY	2329	TTTTTACTCATGTTTCCAAAGAGTGGTGTGCTGTGATGAAGAACTCTTACTTGTGGCTTATA	2388
DB	1748	TTTTTACTCATGTTTCCAAAGAGTGGTGTGCTGTGATGAAGAACTCTTACTTGTGGCTTATA	1807
QY	2389	AAGATGTCAACAAAGCTGTGATGAGGTCAGTACAGGTTTCTATATCTAGTAGCAAGACAG	2448
DB	1808	AAGATGTCAACAAAGCTGTGATGAGGTCAGTACAGGTTTCTATATCTAGTAGCAAGACAG	1867
QY	2449	TAGTACAAATCTGTGGACATAGTTTGGAAACAAAGTCTTACAGATATCTCAGAGATCTTGTG	2508
DB	1868	TAGTACAAATCTGTGGACATAGTTTGGAAACAAAGTCTTACAGATATCTCAGAGATCTTGTG	1927







Db 4088 CCTGGAGCAGGTGTGGTTTACCATCCAGGCAATCGAATACTGTTGGAGATGAGGAA 4147  
Qy 4717 AACCTCTGTTGGAGCACTTCAAAATAGCAGCATTAATGGTCTGAAAGCAATTAATGCACT 4776  
Db 4148 AACCTCTATTTGGAGCACTTCAAAATAGCAGCATAGCGGTCTGAAAGCGCTAAATGCACT 4207  
Qy 4777 TTGCAAGTTGCACAGAGGATTAACCTGTCCTCAGGTCTGATACAGAAACATCTGGTTCGTC 4836  
Db 4208 TTGCAAGTTGCACAGAGGATTAACCTGTCCTCAGGTCTGATACAGAAACATCTGGTTCGTC 4267  
Qy 4837 TTCTATCAGTTGTTCTTCTTAACATAAAATCAGAGATACACCATGCTTCTGCTATAG 4896  
Db 4268 TCCTGTCACTTATCTTCTTAACATGCAATCAGAAATACACAGGCTTCTGCTGTGG 4327  
Qy 4897 ATCTGTTTCATGTTTGGTGGTGTGTTAGCAATTCCTATCCATTCCTGTAATGGAGAC 4956  
Db 4328 ATCTCTTCCATGTTCTGTCGCGCAGTCTTAGCGCTTCCATCTCTGTAATGGAGTACA 4387  
Qy 4957 CTGTTGATCTGAGCTTCTTCACTGTTAGTCTTCTTATTAACCACTTATCTTCCATT 5016  
Db 4388 CCGTGGATCTGAGCGGTGCGCACTAGTCTTCTTATTAACCACTTATCTTCCATC 4447  
Qy 5017 TGATCAACATGGCACATGCTTTCAGATACTTACAGTACACA-----CAG 5064  
Db 4448 TGATCAACATGGCACATGCTTTCAGATACTTACAGTACACA-----CAG 4507  
Qy 5065 GCCTACCTCTGCTCAGGTTCAAGAGACAGTGAAGAGGCTCAATTCGCACTCTCTTCT 5124  
Db 4508 GCGCGCGCTTCTGAGCGGTGAAGAGGATAGTGAAGAGGCTCGCTGTCATCTGCTTCT 4567  
Qy 5125 TTGCAAGAAATTTCTCAATATACAGTGGCTCAATTTGGGTGATATTTCTGCTGTAAT 5184  
Db 4568 TTGCGAAGTGTGCGACACACAGACGCGCTCACTGGGTGGGTGCTCCCGCTGGTACC 4627  
Qy 5185 TGTGGGTCTCACTGAAGATGATCATCCCTTATCTTCTGCTGCTGCTGCTGCTGCTTTC 5244  
Db 4628 TGTGGCTCTCCCTGAGGAAGGCACTACCCCTTACCTCGCTGCTGCTGCTGCTTTC 4687  
Qy 5245 ACTATTTACTTGGGTAACTCGCTCAGGAACTGACATCAATTTCTGCAAGAGGAGT 5304  
Db 4688 ACTATTTACTTGGAGTAGCTCGCTGAGAACTGTTTGGCAATTTCTGCTGAAGAGAT 4747  
Qy 5305 ACAGTCACTCTGTAGCTATCTATCTTACCTAATAATTTGTTCTGCTCTTCCAGGAAT 5364  
Db 4748 TCAGTCACTCTGTAGCTATCTATCTTACCTAATAATTTGTTCTGCTCTTCCAGGAAT 4807  
Qy 5365 ATTGGGATCTGTAAAGGCTTGTCTCAGAGTGTGTGAGATCCTGCTTACTTAACT 5424  
Db 4808 ATTGGGATCCATTAAGGCTTCTAACAAGTGTGTGAGATCCTGCTTACTCAAGT 4867  
Qy 5425 GTTTGAAGCAAAAACACCGTGTGTGAGTACCTTAGAATAAGTAATGTTGATAGAGC 5484  
Db 4868 CTTTGAAGCAGAAAGTGTGTGTGAGTACCTTAGAATAAGTAATGTTGATAGAGC 4927  
Qy 5485 TTCCTGATGATAGTGTGCTTCTGATCAAGCTTCTCAATTCAGGTGCGCAAGTCTG 5544  
Db 4928 TTCCTGAGGACTACAGCTGTCTTCAATCAGGCTTCTCAATTTAGGTGTCCAGGTCTG 4987  
Qy 5545 CAGATGATGCGGAAGCATCTGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5604  
Db 4988 CAGATGATGCGGAAGCATCTGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5047  
Qy 5605 AGAATTTGCTGCCAGCAAAATTTGAAACGGGAGAGGTTGGAGCTTGGCAATTTTTCAG 5664  
Db 5048 AGAATTTGCTGCCAGCAAAATTTGAAACGGGAGAGGTTGGAGCTTGGCAATTTTTCAG 5107  
Qy 5665 CACTTACTGTGAGCCGAGTGTGCTTCTTAAATAATCAGAAATGCGAGTGTGCTC 5724  
Db 5108 CGCTTCAATTTGGTGTGCTGAGTCTGCAATTTCTTAAATAATCAGAAATGCGAGGTGCTC 5167  
Qy 5725 TGGTTGAAGTAAAGCCAGAGCTGTGCTTCTTAAATAATCAGAAATGCGAGTGTGCTC 5784  
Db 5168 TGGTTGAAGTAAAGCCAGAGCTGTGCTTCTTAAATAATCAGAAATGCGAGTGTGCTC 5227

Qy 5785 AAAACAGACCCCTGGCTGAAGAGGGGCAACCCCTTCAATTTATCTCGTGAAGCGGTATCGGA 5844  
Db 5228 AAAACAGACCCAGGGCTAAAGAGAGGAAACCCACTTCAATTTATCTCGGAGCGGTATCGGA 5287  
Qy 5845 AGCTCCATTTGCTCTGGCAACAACACTGCATTTATAGAGAGATGCTAGGAGCCAGAGA 5904  
Db 5288 AGCTGCATTTGCTCTGGCAACAACACTGCATTTATAGAGAGATGCTCGAGCCAGAGA 5347  
Qy 5905 CTAATCAGATGTTATTTGGATTCAACTGGCAGTTACTGTGAGCTCCAACTCTGCTCTCAAG 5964  
Db 5348 CTAATCAGATGTTATTTGGATTCAACTGGCAGTTACTGTGAGCTCTGAGTCTGCTCTCAAG 5407  
Qy 5965 ACAATCAAAATGACGACAGTATGAAAGGCTGATTTCAAAATTTATGAAAACTTTCTGAGG 6024  
Db 5408 ACAATCATGAGTGAATCAATAATAAAGACTGATCTAAATTTCTAGAGAACTTTCTGAGG 5467  
Qy 6025 GCTGGGAAGTATTTGGAGGGTCTTTGCTCCATGTCAGAGTTCACCTTACATCAATAAAT 6084  
Db 5468 ACGGGGGAAGTATTTGGAGGGTCTTTGATCCATGTCAGATTTCAACATTAATAAAT 5527  
Qy 6085 ATTTCTTAATGGAGTATTTGCTTCAATTAGCAAACTATGCTTCAAGGAAAT-AGGAC 6143  
Db 5528 ATTTCTTAATGGAGTATTTGCTTCAATTAGCAAACTATGCTTCAAGGAAAT-AGGAC 5587  
Qy 6144 ATAGATCAATCTGT-----TTT 6160  
Db 5588 ATAGATTAATGTTTATGTTCTAGAACTAAAGAAATGCTTGTTCATCCAGTGTCTAT 5647  
Qy 6161 ATGTGCTAGTATTTCCAGGAATTTATCCCTTCAATTTGCTCATTTTCAATTTTATTT 6220  
Db 5648 TTCTGCTAATATTTCCAGAACTCTTTCCCTTCAATTTGCTCATTTTCAATTTTAT 5707  
Qy 6221 CATCCACTTGTGTAGATGAAGTCACTCAACAGATTGTAGACATTTTATGTTGTGTTAAC 6280  
Db 5708 CACCACCTGTTTATGAGTCAATTAAGCATTTTGTGACATTTCTCATCTGCTAAC 5767  
Qy 6281 TCTTCTGCAATTTTGTATTTTGGTGT 6307  
Db 5768 ATCTCTGCACTTTGTTATTTGGTGT 5794

## RESULT 8

AAC86933

ID AAC86933 standard; cDNA; 6395 BP.

XX AAC86933;

XX AAC86933;

XX 02-APR-2001 (first entry)

XX Nucleotide sequence of Mouse Ubr1 protein.

DE Ubr1; E3-type protein; ubiquitin system; ubiquitin-protein ligase;

XX N-end rule pathway; intracellular pathogen; Lysteria monocytogenes;

XX Yersinia enterocolitica; muscle wasting; infection; ss.

XX Mus sp.

XX Key

XX CDS

XX Location/Qualifiers

XX 115..5388

XX /\*tag= a

XX /product= "Ubr1"

XX US6159732-A:

XX 12-DEC-2000.

XX 11-JAN-1999; 99US-00228317.

XX 02-DEC-1997; 97US-00982956.

XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.

XX

PI Kwon YT, Varshavsky A;  
XX WPI; 2001-090278/10.  
DR P-PSDB; AAB31162.  
XX  
PT Inhibiting the N-end rule pathway in mammalian cells for treating  
PT infections and various diseases associated with muscle tissue wasting, by  
PT inhibiting the expression of Ubri gene.  
XX  
XX Claim 3; Col 15-28; 18pp; English.  
XX  
CC The present sequence encodes a murine Ubri enzyme. Ubri is an E3-type  
CC protein of the ubiquitin system. Specifically, it is a ubiquitin-protein  
CC ligase. The enzyme is specific for destabilising residues exposed at the  
CC N-terminus of protein substrates. Inhibition of the expression of Ubri  
CC gene in a cell results in inhibition of the N-end rule pathway. The  
CC method is used for treatment of mammalian cells infected with an  
CC intracellular pathogen, e.g. *Lysteria monocytogenes* or *Yersinia*  
CC *enterocolitica*. Inhibition of N-end rule pathway is also useful for  
CC treating various diseases associated with wasting of muscle tissue and  
CC infections  
XX  
SQ Sequence 6395 BP; 1802 A; 1376 C; 1519 G; 1698 T; 0 U; 0 Other;  
Query Match 70.3%; Score 4435.4; DB 5; Length 6395;  
Best Local Similarity 86.3%; Pred. No. 0;  
Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;  
QY 590 TTCAGGGCCGTCGTAAGTGTGTCCTCTCTC-TCGACCGGCCACAGGTTTCGCT 648  
DB 8 TTCAGGGCCGTCGTAAGTGTGTCCTCTCTCCTCGCGCGGCCACAGGTTTCGCT 67  
QY 649 TGCCTCTGCGCGGGGTGCGCAACTGACAGGGGTGATTTCCCTCAGATGCGGACGAG 708  
DB 68 AGCTGGCGCGGGGTGCGCAACTGCGGGGTTCGTTCCCTTAAGATGCGGACGAG 127  
QY 709 AGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTACCCAGACCCCTCAGCGTC 768  
DB 128 AGATGGACGGCGCGAGAGATGAGCTGACGCCGGAGCCCTCCCTGCGCCGCGAGCGG 187  
QY 769 TGGCATCTTGTTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGTCATCATTTGGCAC 828  
DB 188 CGGCATCTGTTGGGATCAGCAAGTTGATTTCTATCTGCTTTCTTATCATCATTTGGCAC 247  
QY 829 AATTGGTCCAGAAATTTACTTTGCTGAATGGACCCAGCTTGGAAAGCAGGAGAA 888  
DB 248 AATTAGTCCAGAAATTTATTTGCTGAGATGGACCCAGTTTGGAAAGCAGGAGAGA 307  
QY 889 GTGTACAAATGTCAATATTTCACTCCACTGGATGGTACTTATTTGGAGAGATCCAGATA 948  
DB 308 GTGTACAGATGTAATATCTACTCTCTTTGGAGTGGTACTTATTTGGAGAGATCCGATA 367  
QY 949 TTTGCTTAGAGAAATTTGAAGCAGCTGGAGCAATTTCACTTTGTTGGGAGGTTTCAAAA 1008  
DB 368 TTTGCTTAGAGAAATTTAAACACACAGCTGGAGCGTTCCAGTTGTTGGGAGGTTTCAAAA 427  
QY 1009 GTGGAGAGACAACTATTTCTTGACGGATTTGCAATTTGATTCACATGTTGTA 1068  
DB 428 GTGGAGAGAAACAAATATTTCTGTAGGGATTTGCAATTTGATTCACATGTTGTTGTA 487  
QY 1069 TGGACTGCTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGATCTTCTACTG 1128  
DB 488 TGGACTGCTTCNAAGTAGTGTTCATAAAACCATCGTTTACAAGATGATCTTCTACTG 547  
QY 1129 GAGGAGGGTCTGTGACTGTGGAGACACAGAGGCAATGGAAACTGGCCCTTTTGTGTA 1188  
DB 548 GAGGGGCTCTGTGACTGTGGAGACACAGAGCGTGGAAACTGGCCCTTTTGTGTTGG 607  
QY 1189 ATCATGAACCTGGAGAGAGAGTACTATTAAGAGAAATTCAGCTGTTCGTTGATGAAG 1248  
DB 608 ATCAGAGCTGGAAGAGAGAGTACTACAAAGAGAGCTTACATTTGCCCATTTGAATGAAG 667  
QY 1249 AGGTAATTTGCCAAGCCAGGAAATATTTTCTTTCAGTGTATAAAATATGTCGTAGAAATGA 1308

DB 668 AGGTGATTGCTCAAGCCAGGAGAAATATTCCTTTCGGTGTAAATAATACATTTGTAGAAATGA 727  
QY 1309 CTATATGGGAGAGGAGGAGAAAGAACTGCTCCTCTGAATCCAGATAAGGAGGAGAAATCAAA 1368  
DB 728 CTATATGGGAGGAGGAGGAGAAATTTGCTCCTGAATGCGAGATAAGGAGGAGAAATGAAC 787  
QY 1369 GATACATATGCTGCTCTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGCC 1428  
DB 788 GATACATATGCTGCTCTTTCAACGATGAGCACCATTGCTATGATCATGTGATCTACAGTC 847  
QY 1429 TACAAGAGCTCTTGAATGAGCTCCAGAGGCCAGTTGTCATACCACTGCGCATTTGACA 1488  
DB 848 TGCAGAGAGCTCTAGATTGCGAGCTTCAGAGGACACAGCTGACACAGCACTGCCATCGACA 907  
QY 1489 AAGAGGGTCTGCGGGCTTTAAGCGGAGCTTATGCTGCTGCGGAGGAGGAGGAGGAGG 1548  
DB 908 AAGAGGGTCTGCGGGCTGTCAAGCAGGTGTGTATGCCACTTGCAGAGGAGGAGGAGG 967  
QY 1549 ATATAAAGAGTCAATTCAGAAATGCTCTCAACATCCACTTTCATGTAGAAGTATTAACACT 1608  
DB 968 ATATAAAGAGTCACTCAGAGGACGCTCTCTCAGCACCCCTCCATGTGAGTGTGCACT 1027  
QY 1609 CAGAGATTTATGGCTCATCAGAAATTTGCTTTGCGTCTTGGTTCCTGATGAACAAAATTA 1658  
DB 1028 CCGTGGTTATGGCTTCACAGAAATTCGCTCTGCGCTTGGCTCTCGATGAACAAAATTA 1087  
QY 1659 TGAGCTATTCAGTGACCTTTAGGCGAGCTTTTGGCCAGCATGCTTTAGAGAGCAACTG 1728  
DB 1088 TGAGCTATTCAGTGACCTTTAGACAGATATTTTTGGCAGGCTGCTCGTAGAAGCACTG 1147  
QY 1729 ACTCGAGAGATCCCTGCTCTAAGCAGGTAAATGCTTTTGGGATGCAAGCTTTATTAAG 1788  
DB 1148 GCTCTGAATATCCCTGCTCTTATAGCAGACTAATGCTTTGGGATGCAAACTTTATTAAG 1207  
QY 1789 GTGCCCTTAAGATCCTTCATGAATTTGATCTTACAGAGTTTATTTTATGGAGATGAATACA 1848  
DB 1208 GTGCCCTTAAGATCCTTCATGAATTTGATCTTATAGTAGTTTATTTATGGAGATGAATACA 1267  
QY 1849 AAAAACTTTTGTCTATGGAATTTGTGAAGTATTTAATGAACCTGACAGAGAAATATACA 1908  
DB 1268 AAAAACTTTTGTCTATGGAATTTGTGAAGTATTTAATGAACCTGACAGAGAGTACATCA 1327  
QY 1909 GTGATGATCATGACAGAGATATCTCTATAAATGCTGACATTTTTCAGTTTTCAGATGTTTCT 1968  
DB 1328 GCGAGACCAAGAGAGAGATCTCCATTAACCGCCCTGTCGCTGAGATGCTCACCGTCC 1387  
QY 1969 CTACTCTGGCTGACATCTTATGAAGAGCAGAAATGTTATCTCTGTCATTTACTGAAACTC 2028  
DB 1388 CGACCTTGGCCCGGCATCTTATTCAGAGCAGAGATGTTATTTCTGTCTCATTTACTGAAACGC 1447  
QY 2029 TGCTAGAGTTTTCACCTGAGTACTTGGACAGGACAAATTAATTCACCTCCAGGTTTATA 2088  
DB 1448 TGCTAGAGTTTTCACCTGGAATCTTGGACAGGACAAATTAATTCACCTCCAGGTTTATA 1507  
QY 2089 GCCAGGACAAATTTGGGAGAGTATATGACAGTAAATATGACCTAAAGTATATCTGATCA 2148  
DB 1508 GCCAGGACAAATTTGGGAGAGTCTACGAGTTTATGACCTTAAAGTATATCTGATTA 1567  
QY 2149 GCAAAACCAATATGACAGAAAGATTAAGAATGACAGTTTCTTGAAGTTTTCGATCTT 2208  
DB 1568 GCAAGCTGTCTATTCAGACAGAACGATTAAGAGCGCAGTTCTTGGAGAGTTTTCGGTCTT 1627  
QY 2209 TTTTGAAGTTTTCACCTGATGCGGAGTGGAGAAATCCGAGACAGGTTTGGCAAC 2268  
DB 1628 TTCTGAAGTTTTCACCTGTATGCGGAGTGGAGAAATCAGAGACAGTTTGGCAAC 1687  
QY 2269 ACATTTGAAGTGAATCTGATTTGGGAGGCTGCAATGCTTATACAGATGCAATTTGAAGATA 2328  
DB 1688 ACATTTGAAGTGGACCTGACTGGGAGGCTGCCATGCTATACAGATGCAACTAAAGATA 1747  
QY 2329 TTTTACTCAAGTTTCAAGAGGTGTGCTTGTGATGAAGAACTCTTACTTGTGCTTATA 2388



Qy	4537	AGTCTTCGATTAAATAATTCAAATAGCAATCAAGGAAATGGTTATTCTCTTTGGCCACAACAA	4596
Db	3968	AGTCTTCGGTGAATAATTCAAATAGTATCAAGGAAATGGTCAATCTCTTCGCCACAACAA	4027
Qy	4597	TTTATAGAAATTTGAAATTTGAAAGTGCCACCTGATGAAAGGGATCTCGAGTCCCATGCTGA	4656
Db	4028	TTTACAGAAATTTGGCTTGAAATGGCTCTCTGATGAACCTAGACCCACGAGTGCCCATGATGA	4087
Qy	4657	CCTGGAGCACTCGCGCTTTCACATATCAGGCAATTTGAAATCTATTATTTGGGAGATGAAGGAA	4716
Db	4088	CCTTGGAGCAGTGTGGCTTCCACATCCAGGCATTCGAAACCTGTTGGGAGATGAAGGAA	4147
Qy	4717	AACCTCTGTTTGGAGCATTCAAAATAGGCAGCATAAATGCTCTGAAAGCAATTAATGCAGT	4776
Db	4148	AACCTCTATTGGAGCACTTCAAAATAGACAGCATAGCGGTCTGAAAGCGCTAATGCAGT	4207
Qy	4777	TTGCAGTTGCACAGAGGATTTACCTGTCTCAGGTCTCTGATACAGAAACATCTGTTTCGTC	4836
Db	4208	TTGCAGTTGCACAGAGGCTTACCTGGCTCAGGTCTCTGATACAAACATCTCTGGCTCGGC	4267
Qy	4837	TTCTATCAGTTGTTCTTCTTAAACATAAAATCAGAAGATACCAATGCGCTTCTGCTATATAG	4896
Db	4268	TCTGTCAGTATTCTTCTCTTAACTGCAATCAGAAATACACAGGCTTCTGTCTGTGG	4327
Qy	4897	ATCTGTTTTCATGTTTGGTGGGTCTGTGTTAGCATTTCCATCTTGTATTTGGGATGAACC	4956
Db	4328	ATCTCTTCCATGTTCTGTCGGCGCAGTCTTAGCGTTCCCATCTTGTATTTGGGATGAACA	4387
Qy	4957	CTGTTGATCTGCAGCCCTTCTCAGTTAGTTCTTCTTATAACACAGCTTTATCTCTTCCATT	5016
Db	4388	CCGTGGATCTGCAGCCGCTCGCCACTTAGTTCTTTCATATAACCACTCTATCTCTTCCATC	4447
Qy	5017	TGATCACCATGGGCACATGCTTCAGATATCTACTTACAGTAGACA-----CAG	5064
Db	4448	TGATCACCATGGCGCACATGCTTCAGATATCTCTTCAGATATCTCTTTCACAGATACAGATCTGTCTCCAG	4507
Qy	5065	GCCTTACCCCTTGTCTCAGGTTCAAGAAACAGTGAAGAGGCTCAATTCGCAATCTTCTTTCT	5124
Db	4508	GGCGCGCTTGTCTGAGGGTGAAGAGATAGTGAGGAGGCTCGCTGTGCATCTGCTTTCT	4567
Qy	5125	TTGCAGAAATTTCTCAATATATACAGTGGCTCCATTTGGGTGTGATATTTCTTGGGTGTGTTT	5184
Db	4568	TTTGTGAAGTGTTCGACACACACAGCGCCCTCACTGGGTGCGGTGCTCCCGGTGTGTACC	4627
Qy	5185	TGTGGGTCTCACTGAAGAAATGGCATCACCCCTTATCTTCCGTGTGTGCATTTGTTTTCC	5244
Db	4628	TGTGGGTCTCTCCCTGAGGAACGGCATCACCCCTTACCTCCGCTGTGTGTGCACTGCTTTTCC	4687
Qy	5245	ACTATTTACTTTGGGGTAACTCCGCTCAGGAACTGCATAGCAATTTCTGCAGAAAGGAGAGT	5304
Db	4688	ACTATTTACTTTGGAGTAGCTCCGCTTGAAGATCTGTTTGGCCAAATCTCTGTGAAGGAGAAT	4747
Qy	5305	ACAGTGACTCTGTAGTATCTATCTTTTACCTACAAATTTGTGTTCTGTCTCTTTCAGGAAAT	5364
Db	4748	TCAGTGCACTCTGTAGTATCTATCTTTTACCCCAAAATTTGTTTCTCTGCTTTTTCAGGAAAT	4807
Qy	5365	ATTGGGATCTGTAAAGCCCTTGTCTCAGAGGTGTTGTCAGATCTCTGCCCTTACTTAACT	5424
Db	4808	ATTGGGATACATTAAGGCCCTTACTACAGAGGTGTTGTGAGATCTCTGCCCTTACTCAAGT	4867
Qy	5425	GTTTGAAGCAAAAAAACACCGTGTGCAGGTACCCCTAGAAAAAGAAATAGTTTGTATAGAGC	5484
Db	4868	CTTTGAAGCAGAAAAGTGTGTGTGAGTACCTTAGAAAAGAAATAGTTTGTATAGAGC	4927
Qy	5485	TTCTCTGATGACTATAGCTGTCTCTGAAATCAAGCTTCTCAATTCAGGTGCCACGGTCTG	5544
Db	4928	TTCTCTGAGGACTACAGCTGTCTTCTTAAATCAGGCTTCTCACTTTTAGGTGTCCACGGTCTG	4987
Qy	5545	CAGATGATGAGCGAAGCATCTGTCTCTCGCTTTTCTGTGGGGCTATACTATGTTCTC	5604
Db	4988	CAGATGATGAGCGAAGCATCTGTCTCTCTTCTTTTCTGTGTGTGTGTGTGTGTGTGTCTC	5047
Qy	5605	AGAACTTTGTCTGCCAGGAAATTTGTGAACGGGGAAGAGGTTGGAGCTTGTGAATTTTTCACG	5664

Db	5048	AGAACATCTGTTGCCAAGAAATAGTGAATGGGGAAGAGGTGTGAGCGTGGTTTTTCATG	5107
Qy	5665	CACCTTCACTGTGAGCGGAGTCTGCATTTTCTAAAAATTCAGAGAATGCCGAGTGGTCC	5724
Db	5108	CGCTTCATTTGTGTGTGAGTCTGCATTTTCTTAAATTCGAGAAATCGAGGTGGTCC	5167
Qy	5725	TGTTGAAGGTAAAGCCAGAGCGGTGTGCCCTATCCAGCTCTCTTACTTGGATGAATATGGAG	5784
Db	5168	TGTTGAAGGAAAGCCAGAGCGGTGTGCCCTATCCAGCGCTTACTTGGATGAATATGGAG	5227
Qy	5785	AAACAGACCCTGGCTGTGAGAGGGGCAACCCCTTTCATTTATCTCTGTGAGCGGTATCGGA	5844
Db	5228	AAACAGACCAGGGCTAAAGAGAGGAAACCCACTTTCATTTATCTCGGAGCGGTATCGGA	5887
Qy	5845	AGCTCCATTTGGTCTTGGCAACACACTGCATTAAGAGAGATTGCTAGGAGCAAGAGA	5904
Db	5288	AGCTGCATTTGGTCTTGGCAACAGCACTGCATTAAGAGAGATTGCTCGGAGCAGGAGA	5347
Qy	5905	CTAATCAGATGTTATTGGATTCAACTGGCAGTTACTGTGAGCTCCAACTCTGCCCTCAAG	5964
Db	5348	CTAATCAGATGCTATTGGATTAACTGGCAGTTACTGTGAGCTTCACTTCTGCCCTCAAG	5407
Qy	5965	ACAATCAAAATGACGACAGTAGTAAGGCTGATTCAAAATTAAGAAACTTCTCTGAGG	6024
Db	5408	ACAATCATGAGTGACATCAATAATAAGACTGATCTAAAAATTCAGAGAACTTCTCTGAGG	5467
Qy	6025	GCTGGGAAAGTATTGGAGGGTCTTTTGTCTCATGTCCAGGTTCACATTACATCAATAAAAT	6084
Db	5468	ACGGGGAGATATTGGAGGGTCTTTGATCCATGTCCAGATTCAACATTAAGCTTCAAGGGA	5527
Qy	6085	ATTTCCTTAATGGAGTATGCTTTCAATTAGCAACATATGCTTTCACAGGAAAAA-AGGAC	6143
Db	5528	ATTCCCTTAATGGAAATATTGCTTTCAATTATCAAAACATTAAGCTTCAAGGGA	5587
Qy	6144	ATAGATCAATCTGT-----TTT	6160
Db	5588	ATAGATTAATGTTTATTGTTCTAGAACACTAAAGAAATGCTTGTTCATCCAAGTGTCTAT	5647
Qy	6161	ATGTGCTAGTATTTCAGAGAAATTTATTCCCTTTCATAAATTTGTCTCATTTTCAATTTATTT	6220
Db	5648	TTCTGCTAATATTTCAGAAAATCTCTTCCCTTCATACTGTCTAGTTCAATTCATAT	5707
Qy	6221	CATCCACTTGGTAGATGAAGTCAAGTCAACACAGTTGTAGACATTTATGTGTGGTTAAC	6280
Db	5708	CACCCACTGTTTAATGAGGTTCACATTAAGCAATTTGTGGACATTTCTCCATCTGGCTAAC	5767
Qy	6281	TCCTCTGCAATTTGTAATTTGGTGT	6307
Db	5768	ATCTCTGCACCTTTGTATTGTTGGTGT	5794
RESULT 9			
ADA53348			
ID	ADA53348 standard; cDNA; 3059 BP.		
XX	ADA53348;		
XX	20-NOV-2003 (first entry)		
DT	Human coding sequence, SEQ ID 916.		
XX	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;		
KW	Gene Therapy; human; secretory protein; membrane proteins; cancer;		
KW	inflammatory disease; osteoporosis; neurological disease; gene; 58.		
OS	Homo sapiens.		
XX	EP1293569-A2.		
PN	19-MAR-2003.		
XX	21-MAR-2002; 2002EP-00006586.		
PD			
XX			

	XX	14-SEP-2001; 2001JP-003283B1.
	PR	24-JAN-2002; 2002US-0350435P.
	XX	(HELI-) HELIX RES INST.
	PA	(REAS-) RES ASSOC BIOTECHNOLOGY.
	XX	
	PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
	PT	Yanamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irrie R, Tamechika I;
	PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maehuo Y;
	XX	
	DR	WPI; 2003-395539/38.
	DR	P-PFDB; ADA54987.
	XX	
	XX	New polynucleotides encoding full-length polypeptides, e.g. secretory and/or membrane proteins, useful for developing medicines for diseases in which the gene is involved, or as target molecules for gene therapy.
	PS	Claim 1; SEQ ID NO 916; 205pp; English.
	XX	
	CC	The present invention relates to novel human secretory or membrane proteins (ADA54072-ADA55710) and their coding sequences (ADA52433- ADA54071). The coding sequences are useful in the gene therapy of diseases caused by abnormalities of the proteins, e.g. cancer, inflammatory diseases, osteoporosis or neurological disease.
	XX	
	SQ	Sequence 3059 BP; 969 A; 546 C; 694 G; 850 T; 0 U; 0 Other;
		Query Match      48.4%; Score 3055.8; DB 10; Length 3059; Best Local Similarity    99.9%; Pred. No. 0; Matches 3057; Conservative    0; Mismatches    2; Indels    0; Gaps    0;
	Qy	680 GTCAAGTTTCCCTCAAAGATGCGGCACGAGAGCGCTGGAGGTACTCAGAGGATGAAGAATCACG 739
	Dz	1 GTCAAGTTTCCCTCAAAGATGCGGCACGAGAGCGCTGGAGGTACTCAGAGGATGAAGAATCACG 60
	Qy	740 CGCGAGATTACCAGACCCCCTCAGCGTCGCGCATCTTGCTGGGATCMAGAAATGTAATTT 799
	Dz	61 CGCGAGATTACCAGACCCCCTCAGCGTCGCGCATCTTGCTGGGATCMAGAAATGTAATTT 120
	Qy	800 TTATACTGCTTTCTTGGCATCAATTTGGACAATTCGGTCAGAAAATTTACTTTGCTGGAAT 859
	Dz	121 TTATACTGCTTTCTTGGCATCAATTTGGACAATTCGGTCAGAAAATTTACTTTGCTGGAAT 180
	Qy	860 GGACCCAGACTTGGAAAAAGCAGGAGGAAAGTGTAACAATGTCACATATTCCTCCAAGTGA 919
	Dz	181 GGACCCAGACTTGGAAAAAGCAGGAGGAAAGTGTAACAATGTCACATATTCCTCCAAGTGA 240
	Qy	920 ATGTATCTATTTCGGAGAAGATCAGATATTTGCTTAGAAGAAATGAGACACAGTGGAGC 979
	Dz	241 ATGTATCTATTTCGGAGAAGATCAGATATTTGCTTAGAAGAAATGAGACACAGTGGAGC 300
	Qy	980 ATTTCAGCTTTGTGGAGGGTTTTCAAAGTGGAGAGACACCTATTCTTCAGGGGATG 1039
	Dz	301 ATTTCAGCTTTGTGGAGGGTTTTCAAAGTGGAGAGACACCTATTCTTCAGGGGATG 360
	Qy	1040 TGCAATTTGATCCAACATGTGTACTCTGTATGCCATGCTTCAGGACAGCTGTTCAATAAAA 1099
	Dz	361 TGCAATTTGATCCAACATGTGTACTCTGTATGCCATGCTTCAGGACAGCTGTTCAATAAAA 420
	Qy	1100 TCATCGTTTACAGATGCATCTTCTACTGGAGGAGGGTTCTGTGACTGTGGAGACACAGA 1159
	Dz	421 TCATCGTTTACAGATGCATCTTCTACTGGAGGAGGGTTCTGTGACTGTGGAGACACAGA 480
	Qy	1160 GGCAATGGAAAACTGGCCCCCTTTTGTGTAAATCATGAACCTGGAGAGCAGGTACTATAA 1219
	Dz	481 GGCAATGGAAAACTGGCCCCCTTTTGTGTAAATCATGAACCTGGAGAGCAGGTACTATAA 540
	Qy	1220 AGAGAAATTCAGCTGTGCGTTGGAATGAAGAGGTAAATGTGCCAACGACAGAAAAATTTTCC 1279
	Dz	541 AGAGAAATTCAGCTGTGCGTTGGAATGAAGAGGTAAATGTGCCAACGACAGAAAAATTTTCC 600
	Qy	1280 TTCAGTGTATAAATATGTGCTAGAAATGCACTATATGGGAAGAGGAAAAAGAACTGCCTCC 1339



Db 1681 TGATGAAGAACTCTTACTTGTGGCTTATAAAGAATGTCAAAAGCTGTGATGAGTGCAG 1740  
Qy 2420 TACAGATTTCAATCTAGTAGCAAGACAGTAGTACAATCGTGTGACATAGTTTGGAAAC 2479  
Db 1741 TACCAGTTTCAATCTAGTAGCAAGACAGTAGTACAATCGTGTGACATAGTTTGGAAAC 1800  
Qy 2480 AAAGTCCTACAGAGTATCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGAGCCCT 2539  
Db 1801 AAAGTCCTACAGAGTATCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGAGCCCT 1860  
Qy 2540 TGCTGGCTCTTCATGTACCTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGCGATGAATTTGT 2599  
Db 1861 TGCTGGCTCTTCATGTACCTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGCGATGAATTTGT 1920  
Qy 2600 GTCTTTTGAAGACTTTTCAAGTAGAGGTACTAGTGAATATCTCTTTACCGTTCTGGTGT 2659  
Db 1921 GTCTTTTGAAGACTTTTCAAGTAGAGGTACTAGTGAATATCTCTTTACCGTTCTGGTGT 1980  
Qy 2660 GGTTCGCCAGGTGTGTCTGAGATGTGCGAAGAAATGGAATGCTCTCTTATTAGCCAGGT 2719  
Db 1981 GGTTCGCCAGGTGTGTCTGAGATGTGCGAAGAAATGGAATGCTCTCTTATTAGCCAGGT 2040  
Qy 2720 GTTTTATTACCAAGATGTTAAAGTCAGAGAGAAATGTATGATAAGATATCATCATGCT 2779  
Db 2041 GTTTTATTACCAAGATGTTAAAGTCAGAGAGAAATGTATGATAAGATATCATCATGCT 2100  
Qy 2780 TCAGATTGGTGATCTTTAATAGATGCCAATAAGTTCTTGTACTGTGACTTCAGAGGTA 2839  
Db 2101 TCAGATTGGTGATCTTTAATAGATGCCAATAAGTTCTTGTACTGTGACTTCAGAGGTA 2160  
Qy 2840 TGAATCTGCCAGGCTTTTAAACAGACCATATCTCAAAAGACAGGATTTGATTAAACA 2899  
Db 2161 TGAATCTGCCAGGCTTTTAAACAGACCATATCTCAAAAGACAGGATTTGATTAAACA 2220  
Qy 2900 ATATAATACACTAATAGAGAAATGCTTCAGGTCTCTCATATATTGTGGGTGAGCGTTA 2959  
Db 2221 ATATAATACACTAATAGAGAAATGCTTCAGGTCTCTCATATATTGTGGGTGAGCGTTA 2280  
Qy 2960 TGTACTGTGAGTGGAAATGTGAACAAAGAGAGGTCAATAGAGAGAAATCAATTCACCT 3019  
Db 2281 TGTACTGTGAGTGGAAATGTGAACAAAGAGAGGTCAATAGAGAGAAATCAATTCACCT 2340  
Qy 3020 GCTTTGCATTCGACCCATGCCACACAGTCCATTCGCCAAATTTTACCTGAGAGTGAAGA 3079  
Db 2341 GCTTTGCATTCGACCCATGCCACACAGTCCATTCGCCAAATTTTACCTGAGAGTGAAGA 2400  
Qy 3080 TAATGAAACTGGCTTAGAGATGTCTATAAACAAGTGGCCACATTTTAAAGAACCCAGGTGT 3139  
Db 2401 TAATGAAACTGGCTTAGAGATGTCTATAAACAAGTGGCCACATTTTAAAGAACCCAGGTGT 2460  
Qy 3140 ATCAGGCCATGGAGTTTATGAACCTAAAGATGAATCACTGAAAGACTTCAATATGTACTT 3199  
Db 2461 ATCAGGCCATGGAGTTTATGAACCTAAAGATGAATCACTGAAAGACTTCAATATGTACTT 2520  
Qy 3200 TTATCAATTCCTCAAAACCCAGCATAGCAAGGCTGAACATATGCGAAGAAAGAGAGAA 3259  
Db 2521 TTATCAATTCCTCAAAACCCAGCATAGCAAGGCTGAACATATGCGAAGAAAGAGAGAA 2580  
Qy 3260 ACAAGAAAACAAGATGAAGCATTCGCCGCCACACCATCTCTGAATTTCTGCCCTGCTTT 3319  
Db 2581 ACAAGAAAACAAGATGAAGCATTCGCCGCCACACCATCTCTGAATTTCTGCCCTGCTTT 2640  
Qy 3320 CAGCAAGTGAATTAACCTTCTCAACTGTGATATCATGATGTACATTCCTCAGGACCGTATT 3379  
Db 2641 CAGCAAGTGAATTAACCTTCTCAACTGTGATATCATGATGTACATTCCTCAGGACCGTATT 2700  
Qy 3380 TGAGCGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGCTCCAAATGGCTTT 3439  
Db 2701 TGAGCGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGCTCCAAATGGCTTT 2760  
Qy 3440 TCATATTCCTGGCATTTGGGTTTACTAGAGAGAGAACACAGCTTCAAAAGCTCTCTGAAGA 3499  
Db 2761 TCATATTCCTGGCATTTGGGTTTACTAGAGAGAGAACACAGCTTCAAAAGCTCTCTGAAGA 2820

Qy 3500 AGAAGTAACATTTGACCTTTTATCATTAAGGCTTCAAGATTCGGAAGTTTCAGCCATGAATAT 3559  
Db 2821 AGAAGTAACATTTGACCTTTTATCATTAAGGCTTCAAGATTCGGAAGTTTCAGCCATGAATAT 2880  
Qy 3560 ACAAATCTTTTGGAAAAAAGCTCAAGGAATTTCCCACTTGAAGGCCAGAGGACATGAT 3619  
Db 2881 ACAAATCTTTTGGAAAAAAGCTCAAGGAATTTCCCACTTGAAGGCCAGAGGACATGAT 2940  
Qy 3620 AACGTGATATCTCAGATGTTTGAACACAGTGAAGCGATTAAAGAAAAAATCTTGTTTAAT 3679  
Db 2941 AACGTGATATCTCAGATGTTTGAACACAGTGAAGCGATTAAAGAAAAAATCTTGTTTAAT 3000  
Qy 3680 TGTAGCAACACATCAGATCGGAATCTATTAAAGATGAGATTAATCTCATGATAAG 3738  
Db 3001 TGTAGCAACACATCAGATCGGAATCTATTAAAGATGAGATTAATCTCATGATAAG 3059

## RESULT 10

AAH14878  
ID AAH14878 standard; cDNA; 2550 BP.  
XX  
XX AAH14878;  
XX  
XX 26-JUN-2001 (first entry)  
XX  
XX Human cDNA sequence SEQ ID NO:12731.  
XX  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
XX Homo sapiens.  
XX  
XX EPI074617-A2.  
XX  
XX 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-00116126.  
XX  
XX 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
XX  
XX (HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 12731; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the

CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
xx	
SQ	Sequence 2550 BP; 830 A; 454 C; 558 G; 708 T; 0 U; 0 Other;
	Query Match 40.4%; Score 2546.8; DB 4; Length 2550;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	1189 ATCATGAACCTGGAGAGCAGGTACTATAAAGAGAAATTCACGCTGCTCCGTCGAATCAAG 1248
Db	1 ATCATGAACCTGGAGAGCAGGTACTATAAAGAGAAATTCACGCTGCTCCGTCGAATCAAG 60
Qy	1249 AGGTAATTGTCCAAGCCAGGAAATATTTCTTCAGTGATATAAATATGTCGTGAGAAATGA 1308
Db	61 AGGTAATTGTCCAAGCCAGGAAATATTTCTTCAGTGATATAAATATGTCGTGAGAAATGA 120
Qy	1309 CTATATCGGAAGAGGAAAGAAAGCTGCTCTCTGAACCTCCAGATAAGGAGGAGAAATGAAA 1368
Db	121 CTATATCGGAAGAGGAAAGAAAGCTGCTCTCTGAACCTCCAGATAAGGAGGAGAAATGAAA 180
Qy	1369 GATACTATTGTGTCCTTTTCAAATGATGAACACCAATTCATATGACCACGTCAATATACAGCC 1428
Db	181 GATACTATTGTGTCCTTTTCAAATGATGAACACCAATTCATATGACCACGTCAATATACAGCC 240
Qy	1429 TACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCCAGTTGCAATACACATGCCCAATTGACA 1488
Db	241 TACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCCAGTTGCAATACACATGCCCAATTGACA 300
Qy	1489 AAGAGGCTGCTCGGGCTGTTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAGGAAG 1548
Db	301 AAGAGGCTGCTCGGGCTGTTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAGGAAG 360
Qy	1549 ATATAAGAGTCAITTCAGAAAATGTCTCTCAACATCCACTTCATGTAGAGATATTACACT 1608
Db	361 ATATAAGAGTCAITTCAGAAAATGTCTCTCAACATCCACTTCATGTAGAGATATTACACT 420
Qy	1609 CAGAGATTATGGCTCATCAGAAATTTGCTTTGGTCTGTGTTCTCTGATGAAACAAATTA 1668
Db	421 CAGAGATTATGGCTCATCAGAAATTTGCTTTGGTCTGTGTTCTCTGATGAAACAAATTA 480
Qy	1669 TGAGCTATTCAAGTGACTTTAGCCAGATCTTTTGGCCCAAGCATGCCCTTAGAGAGACCTTG 1728
Db	481 TGAGCTATTCAAGTGACTTTAGCCAGATCTTTTGGCCCAAGCATGCCCTTAGAGAGACCTTG 540
Qy	1729 ACTCGAGAAATCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGACTTTTATAAG 1788
Db	541 ACTCGAGAAATCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGACTTTTATAAG 600
Qy	1789 GTGCCGTAAGATCCTTTCATCAATTGATCTTCAGCAGTTTTTTTTTATGGAGATGGAATACA 1848
Db	601 GTGCCGTAAGATCCTTTCATCAATTGATCTTCAGCAGTTTTTTTTTATGGAGATGGAATACA 660
Qy	1849 AAAAATCTTTGCTATGGAAATTTGTGAGTATTATTAAACAATGCAGAAAGAATATATCA 1908
Db	661 AAAAATCTTTGCTATGGAAATTTGTGAGTATTATTAAACAATGCAGAAAGAATATATCA 720
Qy	1909 GTGATGATCATGACAGAAGTATCTCTATAACTGCACCTTTTCAGTTCAGATGTTTACTGTTTC 1968
Db	721 GTGATGATCATGACAGAAGTATCTCTATAACTGCACCTTTTCAGTTCAGATGTTTACTGTTTC 780
Qy	1969 CTACTCTGGCTCGACATCTTATTGAAGAGCAGAAATGTTATCTCTGTCATTACTGAAACTC 2028
Db	781 CTACTCTGGCTCGACATCTTATTGAAGAGCAGAAATGTTATCTCTGTCATTACTGAAACTC 840
Qy	2029 TGCTAGAAGTTTACTGAGTACTTGGACAGGAACAATAAATTCACATCTTCAGGGTTATA 2088
Db	841 TGCTAGAAGTTTACTGAGTACTTGGACAGGAACAATAAATTCACATCTTCAGGGTTATA 900

QY 3169 ATGAATCACTGAAGACTTCAATATGTACTTTTATCATTACTTCCAAACCCAGCATAGCA 3228  
DB |||||  
QY 1981 ATGAATCACTGAAGACTTCAATATGTACTTTTATCATTACTCCAAACCCAGCATAGCA 2040  
DB |||||  
QY 3229 AGGCTGAACATATGCAGAGAAAGAGGAGAAACCAAGAAACCAAGAGTGAAGCATTTGCCGC 3288  
DB |||||  
QY 2041 AGGCTGAACATATGCAGAGAAAGAGGAGAAACCAAGAAACCAAGAGTGAAGCATTTGCCGC 2100  
DB |||||  
QY 3289 CACACCACTCTCGAATCTGCCCTGCTTTCAGCAAAAGTATTAACCTTCTCAACTGTG 3348  
DB |||||  
QY 2101 CACCACCACTCTCGAATCTGCCCTGCTTTCAGCAAAAGTATTAACCTTCTCAACTGTG 2160  
DB |||||  
QY 3349 ATATCATGATCTCATCTTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCTAACT 3408  
DB |||||  
QY 2161 ATATCATGATCTCATCTTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCTAACT 2220  
DB |||||  
QY 3409 TGTGACCGAAGGAGTGTCCAAATGGCTTTTCATATTTCTGGCATTTGGTTTACTAGAG 3468  
DB |||||  
QY 2221 TGTGACCGAAGGAGTGTCCAAATGGCTTTTCATATTTCTGGCATTTGGTTTACTAGAG 2280  
DB |||||  
QY 3469 AGAAGCAACAGCTTCAAAAAGCTCTGAAGAGTAAATTTGCACTTTTATCATTAAG 3528  
DB |||||  
QY 2281 AGAAGCAACAGCTTCAAAAAGCTCTGAAGAGTAAATTTGCACTTTTATCATTAAG 2340  
DB |||||  
QY 3529 CTTCAAGATTGGGAAGTTTCAGCCATGATATACAAATGCTTTTGGAAAACTCAAAGGAA 3588  
DB |||||  
QY 2341 CTTCAAGATTGGGAAGTTTCAGCCATGATATACAAATGCTTTTGGAAAACTCAAAGGAA 2400  
DB |||||  
QY 3589 TTCCCCAGTTAGAGGCCAGAGACATGATAAAGTGTGATATCTTCAGATGTTTGACACAG 3648  
DB |||||  
QY 2401 TTCCCCAGTTAGAGGCCAGAGACATGATAAAGTGTGATATCTTCAGATGTTTGACACAG 2460  
DB |||||  
QY 3649 TGAAGCGATTAGAGAAAAATCTTTGTTAAATTTGTAGCAACCAATCAGGATCGGAATCTA 3708  
DB |||||  
QY 2461 TGAAGCGTTAAGAGAAAAATCTTTGTTAAATTTGTAGCAACCAATCAGGATCGGAATCTA 2520  
DB |||||  
QY 3709 TTAAGATGATGAGATTACTCATGATTAAG 3738  
DB |||||  
QY 2521 TTAAGATGATGAGATTACTCATGATTAAG 2550  
DB |||||  
RESULT 11  
ACN39798  
ID ACN39798 standard; cDNA; 3776 BP.  
XX  
AC ACN39798;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Tumour-associated antigenic target (TAT) cDNA DNAS25984, SEQ ID NO:4163.  
XX  
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
KW central nervous system cancer; bladder cancer; pancreatic cancer;  
KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
KW chromosome identification; chromosome mapping; gene mapping;  
KW gene therapy; cytostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004030615-A2.  
XX  
PD 15-APR-2004.  
XX  
PP 29-SEP-2003; 2003WO-US028547.  
XX  
PR 02-OCT-2002; 2002US-0414971P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Wu TD, Zhang Z, Zhou Y;

DR WPI; 2004-347921/32.  
DR P-PSDB; ABM81609.  
XX  
FT New tumor-associated antigenic target polypeptides and nucleic acids,  
PT useful in preparing a medicament for treating or detecting a  
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
PT prostate cancer or tumor.  
XX  
PS Claim 1; SEQ ID NO 4163; 7273pp; English.  
XX  
CC The invention relates to human tumour-associated antigenic target (TAT)  
CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
CC overexpressed in cancer tissues compared to normal tissues, and may thus  
CC serve as effective targets for the diagnosis and treatment of cancer in  
CC mammals. The invention also relates to the nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids, useful  
CC antibodies, antagonists, binding molecules and compositions are  
CC for diagnosing or treating a cell proliferative disorder associated with  
CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT nucleic acid of the invention  
XX  
SQ Sequence 3776 BP; 1081 A; 703 C; 750 G; 1242 T; 0 U; 0 Other;  
Query Match 27.2%; Score 1716.4; DB 13; Length 3776;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1717; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4591 CACAAATTTATAGAAATTTGGATTGAAAGTGGCCACCTGATGAAGGGATCTCGAGTCCCCA 4650  
DB 1 CACAAATTTATAGAAATTTGGATTGAAAGTGGCCACCTGATGAAGGGATCTCGAGTCCCCA 60  
QY 4651 TGCAGCTGGAGCACTGGCGCTTTTCACTATCCAGGCAATGAAATCTATTGGGAGATG 4710  
DB 61 TGCAGCTGGAGCACTGGCGCTTTTCACTATCCAGGCAATGAAATCTATTGGGAGATG 120  
QY 4711 AAGGAAAACCTCTGTTTGGAGCACTTCAAAATAGGCGAGCATATGGTCTGAAAGCATTA 4770  
DB 121 AAGGAAAACCTCTGTTTGGAGCACTTCAAAATAGGCGAGCATATGGTCTGAAAGCATTA 180  
QY 4771 TGCAGTTTGCAGTTGCGACAGAGGATTACCTGCTCCTCAGGCTCTGATACAGAAACATCTGG 4830  
DB 181 TGCAGTTTGCAGTTGCGACAGAGGATTACCTGCTCCTCAGGCTCTGATACAGAAACATCTGG 240  
QY 4831 TTCCGCTTTTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACATGCTTCTGT 4890  
DB 241 TTCCGCTTTTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACATGCTTCTGT 300  
QY 4891 CTATAGATCTGTTTTCATGTTTGGTGGTGCTGTGTAGCATTCCTCATCTTGTATTGGG 4950  
DB 301 CTATAGATCTGTTTTCATGTTTGGTGGTGCTGTGTAGCATTCCTCATCTTGTATTGGG 360  
QY 4951 ATGACCTGTTGATCTGCAGCGCTTCTTTCAGTTAGTTCTTCTTATAACCCATTTTATCTCT 5010  
DB 361 ATGACCTGTTGATCTGCAGCGCTTCTTTCAGTTAGTTCTTCTTATAACCCATTTTATCTCT 420  
QY 5011 TCCATTTGATCACCATGGCACACATGCTTTCAGATACTACTTTACAGTAGACACAGCCCTAC 5070  
DB 421 TCCATTTGATCACCATGGCACACATGCTTTCAGATACTACTTTACAGTAGACACAGCCCTAC 480  
QY 5071 CCCTTGTCTCAGGTTCAAGAGACAGATGGAAGGCTCATTCGCGATCTTCTTTTGTGAG 5130  
DB 481 CCCTTGTCTCAGGTTCAAGAGACAGATGGAAGGCTCATTCGCGATCTTCTTTTGTGAG 540

QY 5131 AATATTTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTTGTGGG 5190  
DB 541 AATATTTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTTGTGGG 600  
QY 5191 TCTCACTGAAGAAATGGCACTACCCCTTATCTTGGCTGTGCTGCATTTGTTTCCACTATT 5250  
DB 601 TCTCACTGAAGAAATGGCACTACCCCTTATCTTGGCTGTGCTGCATTTGTTTCCACTATT 660  
QY 5251 TACTTGGGGTAACTCGGCTGAGGAATGTCATACCAATTTCTGCAAGAGAGATGACAGTG 5310  
DB 661 TACTTGGGGTAACTCGGCTGAGGAATGTCATACCAATTTCTGCAAGAGAGATGACAGTG 720  
QY 5311 CACTCTGTAGCTATCTATCTTACCTACAAATTTCTTCTGCTTCTCCAGGAATATTTGGG 5370  
DB 721 CACTCTGTAGCTATCTATCTTACCTACAAATTTCTTCTGCTTCTCCAGGAATATTTGGG 780  
QY 5371 ATACTGTAAAGCCCTTGTCTCAGAGGTGTGTCAGATCTCTGCTTACTAAACTGTTTGA 5430  
DB 781 ATACTGTAAAGCCCTTGTCTCAGAGGTGTGTCAGATCTCTGCTTACTAAACTGTTTGA 840  
QY 5431 AGCAAAAAAACAACCGTGTGAGTACCTTAGAAAAAAGAAATAGTTTGTATAGAGCTTCTCG 5490  
DB 841 AGCAAAAAAACAACCGTGTGAGTACCTTAGAAAAAAGAAATAGTTTGTATAGAGCTTCTCG 900  
QY 5491 ATGACTATAGCTGCTCTCTGATCAAGCTTCTCATTTCAAGTGCCTCCAGGTCTGCAGATG 5550  
DB 901 ATGACTATAGCTGCTCTCTGATCAAGCTTCTCATTTCAAGTGCCTCCAGGTCTGCAGATG 960  
QY 5551 ATGACGAAAGCATCTCTGCTCTGCTTTTCTGTTGGGCTATATCTATGTTCTCAGAACAA 5610  
DB 961 ATGACGAAAGCATCTCTGCTCTGCTTTTCTGTTGGGCTATATCTATGTTCTCAGAACAA 1020  
QY 5611 TTTGCTGCAGGAATTTGAAACGGGGAGAGTTTGGAGCTTGTCAATTTTTCACGCACTTC 5670  
DB 1021 TTTGCTGCAGGAATTTGAAACGGGGAGAGTTTGGAGCTTGTCAATTTTTCACGCACTTC 1080  
QY 5671 ACTGTGAGCCGAGCTCTGCATTTTCTTAAATCAGAGATGCGAGTGTCTGCTGTTG 5730  
DB 1081 ACTGTGAGCCGAGCTCTGCATTTTCTTAAATCAGAGATGCGAGTGTCTGCTGTTG 1140  
QY 5731 AAGTAAAGCCAGAGCTGTGCTATCCAGCTCTTCTTGTGATGAATATGAGAAACAG 5790  
DB 1141 AAGTAAAGCCAGAGCTGTGCTATCCAGCTCTTCTTGTGATGAATATGAGAAACAG 1200  
QY 5791 ACCCTGGCTGAAGAGGGGCAACCCCTTCAATTTATCTCGTGGCGGTATCGGAGCTCC 5850  
DB 1201 ACCCTGGCTGAAGAGGGGCAACCCCTTCAATTTATCTCGTGGCGGTATCGGAGCTCC 1260  
QY 5851 ATTTGGCTGCGCAACAACACTGCATATAGAGAGATTTGCTAGGAGCCAGAGACTAATC 5910  
DB 1261 ATTTGGCTGCGCAACAACACTGCATATAGAGAGATTTGCTAGGAGCCAGAGACTAATC 1320  
QY 5911 AGATGTTATTTGATTTCACTGGCAGTTTACTGTGAGCTCCAACTCTGCTCAAGACATC 5970  
DB 1321 AGATGTTATTTGATTTCACTGGCAGTTTACTGTGAGCTCCAACTCTGCTCAAGACATC 1380  
QY 5971 ACAATGACGACAGTAGTAAAGGCTGATTTCAAAATTTATGAAATCTTCTCAGGGCTGGG 6030  
DB 1381 ACAATGACGACAGTAGTAAAGGCTGATTTCAAAATTTATGAAATCTTCTCAGGGCTGGG 1440  
QY 6031 AAGATTTTGGAGGGTCTTTTGTCTCAGTCCAGGTTTCACTTACATCAATAAATATTTCT 6090  
DB 1441 AAGATTTTGGAGGGTCTTTTGTCTCAGTCCAGGTTTCACTTACATCAATAAATATTTCT 1500  
QY 6091 TAATGAGTATTTGCTTTCAATTAGCAACATATGCTTCAAGGAAAAAGACATAGATC 6150  
DB 1501 TAATGAGTATTTGCTTTCAATTAGCAACATATGCTTCAAGGAAAAAGACATAGATC 1560  
QY 6151 AATCTGTTTATGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGTCTCATTT 6210  
DB 1561 AATCTGTTTATGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGTCTCATTT 1620  
QY 6211 CATTTTATTCATCCACTTGTGTAGTGAAGTCAAGTCAACAGTTGTGTAGACATTTTATGT 6270

DB 1621 CATTTTATTCATCCACTTGTGTAGTGAAGTCAAGTCAACAGTTGTGTAGACATTTTATAT 1680  
QY 6271 GTTGTGAACCTCTTCTGCAATTTTGTATTTGTGTGTTT 6308  
DB 1681 GTTGTGAACCTCTTCTGCAATTTTGTATTTGTGTGTTT 1718  
RESULT 12  
ADRO7317  
ID ADRO7317 standard; cDNA; 3598 BP.  
XX AC ADRO7317;  
XX DT 04-NOV-2004 (first entry)  
XX Full length human cDNA useful for treating neurological disease Seq 823.  
DE gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;  
KW osteoporosis; neurological disease; Alzheimer's disease;  
KW Parkinson's disease; dementia; short memory; cancer;  
KW sense or motor function; emotional reaction; fear response; panic;  
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
KW tranquiliser.  
XX Homo sapiens.  
OS  
XX  
XX EP1447413-A2.  
XX 18-AUG-2004.  
XX 12-FEB-2004; 2004EP-00003145.  
XX 14-FEB-2003; 2003JP-00102207.  
XX 09-MAY-2003; 2003JP-00131452.  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
XX Wakamatsu A, Ishii S, Nagai K, Irie R;  
XX WPI; 2004-583265/57.  
XX P-PSDB; ADRO9273.  
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX Claim 1; SEQ ID NO 823; 2686pp; English.  
XX This invention relates to novel, isolated full length human cDNA  
XX molecules and the encoded proteins thereof. Specifically, it refers to  
XX cDNA clones obtained by an oligo-capping method, where none of these  
XX clones are identical to any known human mRNAs. The present invention  
XX describes an immunoassay to identify agonists and antagonists, as well as  
XX antibodies, antisense molecules and siRNAs that can all be used to bind  
XX to and modulate expression of the cDNA molecules. As such, these  
XX molecules are useful for diagnostic markers or therapeutic targets for  
XX the various diseases or morbid states. In particular, they are useful in  
XX gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
XX disease, Parkinson's disease, dementia, short memory and various cancers,  
XX as well as for maintaining equilibrium of sense or motor function, and  
XX for treating emotional reaction, fear response and panic. Accordingly,  
XX they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
XX cytostatic and tranquiliser activities. This polynucleotide is a full  
XX length human cDNA sequence of the invention. NOTE: This sequence is not  
XX given in the sequence listing of the specification but can be obtained on  
XX CD-ROM from the European Patent Office, Vienna Sub-office.  
SQ Sequence 3598 BP; 1027 A; 669 C; 709 G; 1193 T; 0 U; 0 Other;  
Query Match 24.2%; Score 1529.2; DB 13; Length 3598;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 4775 GTTTCAGTTGCACAGAGATTACCTGCTCCTCAGGTCCTGATACAGAAACATCTGTTTCG 4834
Db 1 GTTTCAGTTGCACAGAGATTACCTGCTCCTCAGGTCCTGATACAGAAACATCTGTTTCG 60
QY 4835 TCTTCTATCAGTTGTTCTTCTTAAACATATAAATCAGAAGATACACATGCTTCTGTTCTAT 4894
Db 61 TCTTCTATCAGTTGTTCTTCTTAAACATATAAATCAGAAGATACACATGCTTCTGTTCTAT 120
QY 4895 AGATCTGTTTCAATGTTTGGTGGGTCGTGTTAGCATTCCTTCCATCTTGTATGGGATGA 4954
Db 121 AGATCTGTTTCAATGTTTGGTGGGTCGTGTTAGCATTCCTTCCATCTTGTATGGGATGA 180
QY 4955 CCTCTGTCATCTGCAGCCTTCTTCAAGTAGTTCTTCTTATACACACCTTTTATCTCTCCA 5014
Db 181 CCTCTGTCATCTGCAGCCTTCTTCAAGTAGTTCTTCTTATACACACCTTTTATCTCTCCA 240
QY 5015 TTTGATCACCATGGCACATGCTTTCAGATCTACTTACAGTAGACACAGGCTTACCCCT 5074
Db 241 TTTGATCACCATGGCACATGCTTTCAGACACTACTTACAGTAGACACAGGCTTACCCCT 300
QY 5075 TGCTCAGGTTCAAGAAGACAGTGAAGGCTCATTTCCGATCTTCTTCTTTGCGAATAT 5134
Db 301 TGCTCAGGTTCAAGAAGACAGTGAAGGCTCATTTCCGATCTTCTTCTTTGCGAATAT 360
QY 5135 TTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTTTGGGTCTC 5194
Db 361 TTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTTTGGGTCTC 420
QY 5195 ACTGAAGATGGCATCACCCCTTATCTTTCGCTGCTGCTCATTTGTTTTCCACTATTTACT 5254
Db 421 ACTGAAGATGGCATCACCCCTTATCTTTCGCTGCTGCTCATTTGTTTTCCACTATTTACT 480
QY 5255 TGGGTAACTCCGCTGAGGAATCGATACCAATTTCTGAGAGAGAGATGACAGTGCAT 5314
Db 481 TGGGTAACTCCGCTGAGGAATCGATACCAATTTCTGAGAGAGAGATGACAGTGCAT 540
QY 5315 CTGTAGCTATCTATCTTTACCTACCAATTTGTTCTGCTCTTCCAGGAATATTTGGATAC 5374
Db 541 CTGTAGCTATCTATCTTTACCTACCAATTTGTTCTGCTCTTCCAGGAATATTTGGATAC 600
QY 5375 TGTAAGGCCCTTGCTCCAGAGGTGCTGTCAGATCTGCTTACTTAAACTTGTTGAAGCA 5434
Db 601 TGTAAGGCCCTTGCTCCAGAGGTGCTGTCAGATCTGCTTACTTAAACTTGTTGAAGCA 660
QY 5435 AAAAAACACCGTGTGTCAGTACCTTAGAAAAAAGAAATAGTTTGTAGAGCTTCTCTGATGA 5494
Db 661 AAAAAACACCGTGTGTCAGTACCTTAGAAAAAAGAAATAGTTTGTAGAGCTTCTCTGATGA 720
QY 5495 CTATAGCTGCCCTCCGTAATCAAGCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGA 5554
Db 721 CTATAGCTGCCCTCCGTAATCAAGCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGA 780
QY 5555 GCGAAGCATCTCTGCTCTGCTGCTTTTCTGTGGGCTATATCTATGTTCTCAGAAATTTG 5614
Db 781 GCGAAGCATCTCTGCTCTGCTGCTTTTCTGTGGGCTATATCTATGTTCTCAGAAATTTG 840
QY 5615 CTGCCAGGAATTTGTAACGGGGAAGAGTTGGAAGCTTGCAATTTTTCAGCACTTCACTG 5674
Db 841 CTGCCAGGAATTTGTAACGGGGAAGAGTTGGAAGCTTGCAATTTTTCAGCACTTCACTG 900
QY 5675 TGGAGCCGAGTCTGCTATTTTCTTAAATCAGAGATGCCAGTGGTCTGTTGAAGG 5734
Db 901 TGGAGCCGAGTCTGCTATTTTCTTAAATCAGAGATGCCAGTGGTCTGTTGAAGG 960
QY 5735 TAAAGCCAGAGCTGTGCTATCTATCTCAGCTCTTACTTGGATGAATATGAGAAACAGACCC 5794
Db 961 TAAAGCCAGAGCTGTGCTATCTCAGCTCTTACTTGGATGAATATGAGAAACAGACCC 1020
QY 5795 TGGCTGTAAGAGGGGCAACCCCTTCAATTTATCTCGTGAAGGGTATCGAAGCTCCATTT 5854
Db 1021 TGGCTGTAAGAGGGGCAACCCCTTCAATTTATCTCGTGAAGGGTATCGAAGCTCCATTT 1080
```

```
QY 5855 GGTCTGCCAACAAACATCTGCATTTATAGAAGAGATTGCTAGGAGCCAGAGACTAATCAGAT 5914
Db 1081 GGTCTGCCAACAAACATCTGCATTTATAGAAGAGATTGCTAGGAGCCAGAGACTAATCAGAT 1140
QY 5915 GTTATTTGGATTCAACTGGGAGTTACTGTGAGCTTCCAACTCTGCTCAAGCAATCACA 5974
Db 1141 GTTATTTGGATTCAACTGGGAGTTACTGTGAGCTTCCAACTCTGCTCAAGCAATCACA 1200
QY 5975 ATGACGACAGTAGTAAAGGCTGATTCAAAAATTATGGAAGAACTTTCTGAGGCTCGGAAAG 6034
Db 1201 ATGACGACAGTAGTAAAGGCTGATTCAAAAATTATGGAAGAACTTTCTGAGGCTCGGAAAG 1260
QY 6035 TATTGGAGGCTCTTTTGTCTCCATGTCAGGTTCACTTACATCAATAAATAATTTCTTAAT 6094
Db 1261 TATTGGAGGCTCTTTTGTCTCCATGTCAGGTTCACTTACATCAATAAATAATTTCTTAAT 1320
QY 6095 GGAGTATTTGCTTTCAATTTAGCAAAACATATGCTTCAAGGAAAGGACATAGATCAATC 6154
Db 1321 GGAGTATTTGCTTTCAATTTAGCAAAACATATGCTTCAAGGAAAGGACATAGATCAATC 1380
QY 6155 TGTTTTATGTGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGTCTCATTTCAAT 6214
Db 1381 TGTTTTATGTGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGTCTCATTTCAAT 1440
QY 6215 TTATTTCACTCAGCTTGTGTAGTAGAAGTCACTCAAGCAAGTTGTAGACATTTTATGTGTTG 6274
Db 1441 TTATTTCACTCAGCTTGTGTAGTAGAAGTCACTCAAGCAAGTTGTAGACATTTTATGTGTTG 1500
QY 6275 GTTAACTCTTCTGCAATTTTGTATTTGGTGTGTTT 6308
Db 1501 GTTAACTCTTCTGCAATTTTGTATTTGGTGTGTTT 1534

RESULT 13
ADJ95443
ID ADJ95443 standard; cdna; 6300 BP.
XX
AC ADJ95443;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human cdna encoding Ubiquitin ligase E3alpha II.
XX
KW Human; ss; gene; ubiquitin ligase; E3alpha II;
KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;
KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;
KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;
KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;
KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 295..5562
FT /*tag= a
FT /product= "Ubiquitin ligase E3alpha II"
XX
PN US6706505-B1.
XX
PD 16-MAR-2004.
XX
PP 28-NOV-2000; 2000US-00724126.
XX
PR 08-MAR-2000; 2000US-0187911P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Han H, Kwak K;
XX
XX WPI, 2004-236723/22.
DR P-PSDB; ADJ95444.
XX
PT New nucleic acid molecule, useful for preparing a composition for
```



PT diagnosing, treating or preventing diseases associated with human  
PT E3approximatielya polypeptide, e.g., muscle atrophy.

XX Example 2; SEQ ID NO 3; 104pp; English.

XX The invention relates to a new isolated nucleic acid molecule appearing  
CC as ADJ95441(or its complement) encoding a ubiquitin ligase E3alpha I  
CC protein appearing as ADJ95442. Also included are a vector comprising the  
CC nucleic acid, a host cell comprising the vector, a process of producing a  
CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the  
CC nucleic acid molecule, a reagent comprising a detectably labelled  
CC nucleotide, and a method for determining the presence of a human E3alpha  
CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid  
CC molecule is useful for preparing a composition for diagnosing, treating  
CC or preventing diseases associated with human E3alpha I polypeptide, e.g.  
CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal  
CC cachexia, inflammatory cachexia, muscle wasting disorders associated with  
CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,  
CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting  
CC syndrome. The present sequence encodes human E3alpha II.

XX Sequence 6300 BP; 1842 A; 1235 C; 1478 G; 1745 T; 0 U; 0 Other;

Query Match 21.9%; Score 1381.8; DB 12; Length 6300;

Best Local Similarity 56.5%; Pred. No. 0;

Matches 2947; Conservative 0; Mismatches 2167; Indels 105; Gaps 16;

QY	780	TGGGATCAGCAAGTGTGATTTTATATCTGCTTTCTTCGATCATTTTGGCACAATTTGGTGCCA	839
DB	376	TGGCTGCAAGCACTGACCTCACTAGAGAGTGTACCGATTTAGCCCACTATGTACCC	435
QY	840	GAATTTTACTTTGCTGAAATGACCAGACTTTGGAAAAGCAGGAGGAAGTGTACAAATG	899
DB	436	AAATCTACTGCGGGGTCCCAACCTTTTCCACAGAAAGACATGCTGGCAGCAT	495
QY	900	TCAATTTACTCCACTGGAATGGTACTTATTTGGAGAAGATCCAGATATTGCTTAGAG	959
DB	496	GTTTTGTGGGACCAATGGAATGGTACCTTTGTGTGGAAGATCTCTGCATTTGGATTTCCA	555
QY	960	AAATTTGAAGCAC--AGTGGACATTTGAGCTTTGGGAGGGTTTCAAAGTCGAGAG	1016
DB	556	AACTTGGACAGCAACAACTTCTCATCTTTGTGTGCTGTGTTTTTAAGTAGAGAG	615
QY	1017	ACAACTTATTTGTCAGGGAATGTGCAATGTATGATCCAAATGATGTACTCTGTATGGACTGC	1076
DB	616	CCTACATATTTCTGCAGAGACTGTGCAGTGTATGATCAACTTGTGTTTTGTGCAATGAGTGC	675
QY	1077	TTCAGGACAGTGTTCATAAATCATCTGTTTCAAGATGCATCTTCTACTGGAGAGGG	1136
DB	676	TTTTTGGGAAGTATTTACAGAGATCATCGATATAGGATGACAAATCAGGAGGTGGAGGT	735
QY	1137	TTCTGTGACTGTGGAGACACAGGACATGGAATCTGGCCCTTTTGTGTAAATCATGAA	1196
DB	736	TTCTGTGACTGTGGTGTACTGAGCCCTGGAAAAGGGTCTTACTGTCAAAAACATGAA	795
QY	1197	CCTGGAAGGACGAGGTACT---ATAAAGAGAAATTCACGCTGTCCGTGTTGAATGAAGAGTA	1253
DB	796	CTTAACACCTCTGAAATTTGAGGAAGAGAGGATCTCTTGTTCATTTATCAGAAGATGTG	855
QY	1254	ATTGTCCAGCAGGAAATATTTCTTCAGTGATATAAATATGTCTGTAGAATGACTATA	1313
DB	856	ATAGCAAGAACTTATAACATTTTGTCTATTACGTTTTCGTTATGCAATGAGATAATTAACC	915
QY	1314	TGGGAAGGAAAGAAAGACTGCTCTCGAATCCAGATGAAGGAGGAAATGAAGATAC	1373
DB	916	TGGGAAAGAAAGTGAATTTGCCAGAGATTTAGAGATGGTAGAAGAGTGACACCTAC	975
QY	1374	TATTGTGCTTTTCAATGATGAACACCAATTCATATGACCAGCTCATATACAGCCCTACAA	1433
DB	976	TATTGCACTGCTTTTAATGATGAGGTTTACACCTATGAAAGTATTTTATCTCTTACG	1035
QY	1434	AGAGCTTGTACTGTGAGCTGCGAGGCGCCAGTTGTCATACACTGCCATTTGACAAAGAG	1493

DB	1036	AAAGCTGTTAACTGTACACAAAAGAACCTATTGGTTTTGCAACTACAGTAGATCGAGAT	1095
QY	1494	GGTCTCGGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGAGATATA	1553
DB	1096	GGGCGTAGGCTCTGTTTCGATATGGAGATTTTCAGTATTTGTGAGCAAGCAAAATCAGTAAT	1155
QY	1554	AAGAGTCATTTCAGAAAATGCTCTCAACATCCACTTCATGTAGAATATTACACTCAGAG	1613
DB	1156	GTGGAATATACCACTAGACAGAC---AAAGCCACTCAAAGTTTCAAGTTATGCAATTCGTCT	1212
QY	1614	ATTATGGCTCATCAGAAATTTGCTTGGCTCTTGGTCTCTGATCAACAAAATTAATGAGC	1673
DB	1213	ATTGTGCAATCAGAAATTTGGTTTGAACCTTTTGTCTTGGCTGGGAAGTATTATTGA	1272
QY	1674	TATTCAAGTGAATTTTAGGCAGATCTTTTGGCCAGCATGCTTTAGAGAAGAACCTGACTCG	1733
DB	1273	TATTCAAGTGGCTTTCGCGGATTTTATGTCAAGTTGGTTTACAAAGAGGCGCAGATGCT	1332
QY	1734	GAGATCCCTGCTCATTAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCC	1793
DB	1333	GAATACTCTTCTCTAGTGGACAGACTGATGCTTAGTGTCCAAATTTATGGAAGGTGCT	1392
QY	1794	CGTAAGATCTTTCATGAATGATCTTCAGCAGTCTTTTATGGAGATGCAATACAAAAA	1853
DB	1393	AGGAGTGATATCATCAGTTGTTTATGAGCAGCTGCTGTTATGGATTGGAATACAGAA	1452
QY	1854	CTCTTGTCTGTAATTTTGAAGTATTATTAACAACACTGCAGAAAGAAATATATCAGTGT	1913
DB	1453	CTATTGCTGTTTCGATTTGCAAAAAATTTACCAGCAGTTTGCAGAGAGATTTTATGGAGAT	1512
QY	1914	GATCATGACAGAGTATCTCTATACTGCACTTTTCAGTTTCAGATGTTTACTGTTCTACT	1973
DB	1513	GATCAGCAGCAGAGTGTGCTGCTCTATCTGTCCAGTCTTCCACGCACTACT	1572
QY	1974	CTGCTGCTGACATCTTATGGAAGCAGATGTTATCTCTGTCATTACTGAACTCTGCTA	2033
DB	1573	CTGGCTCGAATGCTCATCACAGAGAAATCTTGATGAGCATTTATCATTAAGACTTTTATG	1632
QY	2034	GAAGTTTT-----ACCTGAGTACTTGACAGGAAACAAATAAATTCACCTTCCAGGTTAT	2087
DB	1633	GATCATTTGACACATCGAGATGCCCGGCGAGATTTTCAGTTTGAAGGATACACTGCTTTA	1692
QY	2088	AGCCAGGACAAATTTGGGAAGAGATATGCGATATAATGTGACCTTAAAGATATCTCGATC	2147
DB	1693	CAAGCTTTCAAATTTAGGAGAGTACAGAGCCTTATTTTAGATCTCAAGATATGTTTAAT	1752
QY	2148	AGCAACCCCAATATGACAGAAAGATTAGAATGCAATGCTTCTTGAAGTTTTCGATCT	2207
DB	1753	AGCAACCCCAATGCTGATGCTGAGCTGAGGAGGAGGTTTCTAGAAAGGTTTGTATGCC	1812
QY	2208	TTTTTGAAGATTTCTTACTGTATGCGGGAATGGAAGAAATCCGAAAGACAGTTGGGCA	2267
DB	1813	TTTTTGGATTTACTTAAATGTATGCGAGGAATGATCCAAATTAACGTCAGTAGGACAA	1872
QY	2268	CACATTAAGTGGATCTTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAT	2327
DB	1873	CATATTGAATGGAACACAGAGTGGGAGCAGCCTTCACTACAAATGAAATTAACACAT	1932
QY	2328	ATTTTACTCATGTTTCCAGAGTGGTGTGCTGATGAGAGAACTTCTTACTTGTGGCTTAT	2387
DB	1933	GTCAATTTCAATGATGCGAGGACTGGTGTGCTTTCAGATGAAAAGTGTTAATCGAAGCTTAC	1992
QY	2388	AAAGAAATGTCACAAAGCTGTGATGAGGTGAGTACCAGTTTTCATATCTAGTAGCAAGACA	2447
DB	1993	AAGAAATGCTCGCTGTACTGATGTCAGTGTCTATGTTGTTTATCTGATGGTGAACAGCA	2052
QY	2448	GT---AGTCAATCGTGTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTGAGAT	2504
DB	2053	ATCACACTAAGCAATTTTGTGGACATTCAGTGGAAACTATCAGATACTGTGTTTCCCAAGAA	2112
QY	2505	CTTGTAAGCATACATCTGCCCACTCTCTAGGACCCCTTGTCTGTCTTCTATGTTTAAAGC	2564
DB	2113	AAAGTTAGCATTCACCTCCCAAGTTTCTCGCTTACTTTCAGGTTTACATGTATTATTAAAGC	2172



Db 4312 ATCCAAAGCATAGAAAGATTTTGGATGATGAAGATAAACCATTTGTTGGTCTTTACCT 4371  
QY 4740 AATAGCAGCATAAATGGTCTGAAGCATTAATTCAGATTTTGCAGTTGTCACAGAGGATACC 4799  
Db 4372 TGCAGACTGGATGACTGTCTTAGGTCTATTGACAGATTTTGGCCGACGACACTGGACAGTG 4431  
QY 4800 TGTCTCAGGTCTGATACAGAAACATCTGGTCTGTTCTATCAGTTGTTCTTCTTCAAC 4859  
Db 4432 GCATCAGTTTTCAGTGGTGCAGGACATTTTGTAACTTTTTCATCACTGGTGGCTAAT 4491  
QY 4860 ATAAATCAGAAATACACATGCTTCTGTCTATAGATCTGTTTCATGTTTGGTGGCT 4919  
Db 4492 GACAGCATGAGAACTTCATGCAATATTAGATATTGACATGTTTCATTTATTTGGTGGC 4551  
QY 4920 GCTGTGTTAGCATTCCTGTTATTTGGGATGACCTGTTGATCTGCAGCTTCTTCA 4979  
Db 4552 TTGGTGTGCTGCTTCTGCTGGTGCAGTG-----TCAGGATTTTTCAGGG 4596  
QY 4980 GTTAGTCTTCTATTAACACCTTTTATCTCTTCCATTTGATCAGCATGGGACACATGCTT 5039  
Db 4597 ATCAGCTTTGGCAGTGGAGACCTTCACATTTTCCATCTGGTTACTATGGCACATCATATA 4656  
QY 5040 CAGATACTACTTACAGTACAGACAGGCC---TACCCTTGTCTCAGGTTTCAAGAACAGT 5096  
Db 4657 CAGATCTTACTTACCTCATGTACAGAGAGAAATGGCATGGATCAAGAAATCCCTCTGT 4716  
QY 5097 GAAGAGGCTCATTCGCGCATCTTCTTTTTCAGAAATTTCTCAATATACAGAGTGGCTCC 5156  
Db 4717 GAAGAAGATCAGCAGTCTTCTGTTGTATTAACACATTCACCATGATATACGGGAGTGCC 4776  
QY 5157 ATGGGTGTGATATTCTGGCTGGTATTTGTGGGTCTCACTGAAGATGGCATACCCCT 5216  
Db 4777 TTGAAGAAATACCATTCCGCTGGCATCTGTGGAGGAGTGTACAGAGCTGGAATCATGCT 4836  
QY 5217 TATCTTCGCTGCTGCTGCTGTTTTCACATTTTACTTGGGTAACCTCCGCTCAGGAA 5276  
Db 4837 TTCTGGAAGTGTCTGCTTTATTTTTCATTACTTAATAGGAGTCTCTTCCCAACCCGAC 4896  
QY 5277 CTGCATACCAATCTCTGAGAAGGAGATGACAGTGCATCTGTAGCTATCTATCTTTACT 5336  
Db 4897 ATTCA---AGTTCCTGGACAGCCATTTTGAACATTTATGTAGCTATCTTCCCTACCA 4953  
QY 5337 ACAATTTGTTCTCTCTTCAGGAATATTCGGATATCTGTAAGCCCTTGTCTCAGAGG 5396  
Db 4954 AACCACTCATTTGCTTTTCAAGAAATAGTGAGATAATGAATTCATGATGAAAGT 5013  
QY 5397 TGGTGTGCAGATCTGCTTACTTAACTGTTTGAAGCAAAACACCGTGTGCTAGGTAC 5456  
Db 5014 TGGTGCCTGTAACAGTGAAGTAAAGATATCTAGAAGGTGAAGAGATGCTATAGATAT 5073  
QY 5457 CCTAGAAAAGAAATAGTTTGTATAGAGCTTCTGTATGATATAGCTGCTCTGATCAA 5516  
Db 5074 CCAAGAGATCTAACAAATTAATAACCTTCCAGAGATTACAGAGCTCTATTAATCAA 5133  
QY 5517 GCTTCTCATTTTCAGGTGCCAGGTCTGCAGATGATGACGAAAGCATCTGTCTCTGTC 5576  
Db 5134 GCATCCAAATTTCTGTGCGCGAAATCAGGTGTGATTAAGACAGAGCCCAACTCTGTGC 5193  
QY 5577 CTTTCTGTGGGCTATATCTATGTTCTCAGAAATTTGTCGACGAAATTTGTGAACGGG 5636  
Db 5194 CTTGTGTGGGATCTCTGCTGTGCTCCAGAGTTACTGCTGCCAGACTGAATGGAAGGG 5253  
QY 5637 GAAGAGTTGGAGCTTGCATTTTTCAGGCATTTCTGTTGGAGCGGAGTCTGCATTTTC 5696  
Db 5254 GAGGATGTAGAGCTGTCACAGCTCACACTACTCTGTGGCTCTGAGTGGGCACTTTC 5313  
QY 5697 CTAATAATCAGAAATGCGAGTGTCTGTTGAAGGTAAAGCCAGAGGCTGTGCTCAT 5756  
Db 5314 CTGAGAGTACGGGAATGTGAGTGTCTATTTTGTAGCTGGCAAAACCAAGGCTGTTTAT 5373  
QY 5757 CAGCTCTTACTTGTGATGAATATGAGAAACAGACCTTGGCTGGAAGGGGCAACCC 5816

Db 5374 TCTCTCTCTTACCTTGTGATGACTATGGGAGAGACGACGAGGACTCAGACGGGGAATCCT 5433  
QY 5817 CTTCAATTTATCTGTGAGCGGTATCGGAAGCTCCATTTGGTGTGCAACAACTGCAATT 5876  
Db 5434 TTACATTTATGCAAGAGCATTTCAAGAAGATTCAGAAGCTCTGGCACCACACAGTGTG 5493  
QY 5877 ATAGNAGAGATTGCTAGGAGCCAGAGACTAATCAGATGTTATTTGGATTCACTGGCA 5935  
Db 5494 ACAGAGGAATTTGGACATGCAAGGAAGCCAAATCAGACTGTTGGCAATTGACTGGCA 5552

RESULT 14  
ADS86865  
ID ADS86865 standard; cDNA; 6300 BP.  
XX  
AC ADS86865;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
DE cDNA encoding human E3alpha ubiquitin ligase, hE3alphaII.  
XX  
KW Human; E3alpha ubiquitin ligase; huE3alphaII;  
KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;  
KW fasting; metabolic acidosis; muscle degeneration; kidney failure;  
KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;  
KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;  
KW inflammatory cachexia; hyperthyroidism; denervation atrophy;  
KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;  
KW gene therapy; ss; gene.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 5'UTR 1..294  
FT CDS /\*tag= a  
FT 295..5562  
FT /\*tag= b  
FT /\*product= "huE3alphaII"  
FT 5563..6300  
FT /\*tag= c  
XX  
US2004185037-A1.  
XX  
PD 23-SEP-2004.  
XX  
PF 15-JAN-2004; 2004US-00758672.  
XX  
PR 08-MAR-2000; 2000US-0187911P.  
XX 28-NOV-2000; 2000US-00724126.  
XX (HANH/) HAN H.  
XX (KWAK/) KWAK K.  
XX Han H, Kwak K;  
XX  
XX WPI; 2004-707854/69.  
XX P-PSDB; ADS86866.  
XX  
XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule  
XX useful for treating and/or preventing renal cachexia or inflammatory  
XX cachexia.  
XX  
XX Claim 1; SEQ ID NO 3; 115pp; English.  
XX  
XX The present invention relates to new orthologue of human E3alpha  
XX ubiquitin ligase, huE3alphaII and huE3alphaII. Most intracellular proteins  
XX are degraded through the ubiquitin-proteasome pathway. Proteins are  
XX marked for proteasomal degradation by conjugation of ubiquitin to the  
XX protein. Conjugation of the ubiquitin molecule involves the activation by  
XX E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,  
XX and then interacts with a specific E3 ligase family member. E3 ligase  
XX binds to proteins targeted for degradation and catalyses the transfer of  
XX ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase

CC determines the specificity of the system. The E3alpha family is the main  
 CC family of intracellular ligases and is involved in the N-end rule pathway  
 CC of protein degradation. E3alpha enzyme binds directly to the primary  
 CC destabilising N-terminal amino acid and catalyses ubiquitin conjugation  
 CC thereby targeting the protein for degradation. The human E3alpha gene is  
 CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-  
 CC proteasome pathway has been determined to be a major cause of rapid  
 CC muscle wasting including, fasting, metabolic acidosis, muscle  
 CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,  
 CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen  
 CC balance, burns, Cushing's syndrome, inflammatory cachexia,  
 CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-  
 CC protein malnutrition. E3alpha plays a role in the overall increase in  
 CC ubiquitination that is associated with and may mediate muscle atrophy in  
 CC catabolic and other disease states. Treatment may be administered by gene  
 CC therapy, cell therapy and antisense therapy methods. The present sequence  
 CC is cDNA encoding human E3alpha1.  
 XX

Sequence 6300 BP; 1842 A; 1235 C; 1478 G; 1745 T; 0 U; 0 Other;

Query Match 21.9%; Score 1381.8; DB 13; Length 6300;  
 Best Local Similarity 56.5%; Pred. No. 0;  
 Matches 2947; Conservative 0; Mismatches 2167; Indels 105; Gaps 16;

QY	780	TGGGATCAGCAAGTTGATTTTATCTGCTTCTTGGCATCTTTGGCACAATTGGTGCCA	839
DB	376	TGGCTGCAAGCACTGACCTCACTAGAGAGTGTTACGACCTTTAGCCCACTATGTACCC	435
QY	840	GAATATTTACTTTGCTGNAATGGACCCAGACCTTGGAAAGCAGGAGGAAGTGTACAAATG	899
DB	436	AAATCTACTGCGGGGTCCTCAACCTTTTCCACAGAAAGACATGCTGGCACAGCAT	495
QY	900	TCATATTTCACTCCACTGGATGGTACTTATTTGCAAGAGATCCAGATATTTGCTTAGAG	959
DB	496	GTTTTGTGGGCAATGGAATGGTACCTTTGTGTGAAGATCTCGCATTTGGATTTCCA	555
QY	960	AAATTTGAAGCAC---AGTGGACATTTTCAGCTTTGTGGGAGGTTTCAAAAGTGGAGAG	1016
DB	556	AACTTGAGCAAGCAACCAACCTTCTCATCTTTGTGTGCTGTTTTTAAGTAGGAGAG	615
QY	1017	ACAACCTATTTTGCAGGAGTTGCAATTTGATCCAAATGATGTACTCTGTATGGACTGC	1076
DB	616	CCTACATATTTCTGCAGAGACTGTGCAGTTGATCCAACTTGTGTTTGTGCATGGAGTGC	675
QY	1077	TTCCAGCAGAGTTGTCATAAAATCATGTTTCAAGATGCATCTTCTACTGGAGAGGG	1136
DB	676	TTTTTGGGAAGTATTTACAGAGATCATCGATATAGGATGACAAATCAGGAGGTGGAGGT	735
QY	1137	TTCTGTGACTGTGGAGACACAGAGCATGGAAATCTGGCCCTTTTGTGTAAATCATGAA	1196
DB	736	TTCTGTGACTGTGGTGTACTGAGCCTGGAAAGAGGGTCTTACTGTCAAAAACATGAA	795
QY	1197	CCTGGAAGCAGGAGTACT---ATAAAGAGAAATTCACGCTGTCCTGTTGAATGAAGAGTA	1253
DB	796	CTTAACACCTCTGAAATTTGAGGAAGAGAGGATCTCTTGTTCATTTATCAGAGATGTG	855
QY	1254	ATTGTCACAGCAGGAAATATTTCTTCAGTGATATAAATATGTCGTAGAAATGACTATA	1313
DB	856	ATAGCAAGAACTTATAAATTTTGTCTATTACGTTTCGGTATGTCAGTAGAATAATTAAAC	915
QY	1314	TGGGAAGAGGAAAGAACTGCTCTGAACTCCAGATAAGGGAGAAATAAGAAAGATAC	1373
DB	916	TGGGAAGAAAGAAAGTGAATTTGCCAGAGATTTAGAGATGTTAGAGAGTGACACCTAC	975
QY	1374	TATTTGTGCTTTTCAATGATGAACACCATTTCAATATGACCAAGTTCATATACAGCCTCAA	1433
DB	976	TATTTGCATGCTGTTTAAATGATGAGTTTCAACCTATGAACAAGTTATTTATATCTCTCAG	1035
QY	1434	AGAGCTCTTGACTGTAGCTCCAGAGGCCAGTTGTCATACCACTGCCATTTGACAAAGAG	1493
DB	1036	AAAGCTGTTAACGTACACAAAGAAAGAGCTATTTGGTTTTTGGCACTACAGTAGATCAGAT	1095
QY	1494	GTCGTGCGGCTGTTAAAGCGGAGCTTATGCTGCTTGGCAGGAAGCAAGGAAGATATA	1553

DB	1096	GGCGTAGGTCGTGTTTCATATGAGATTTTTCAGTATTTGTCAGCAAGCAAAATCAGTAAT	1155
QY	1554	AAGAGTCATTTCAGAAATGTCCTCAACATCCCACTTCATCTAGAGATTTTACACTCAGAG	1613
DB	1156	GTGAGAAATACCAAGTAGACAGAC---AAAGCCACTCAAGTTTCAAGTTATGATTCGTCT	1212
QY	1614	ATTATGCTCATCAGAAATTTGCTTTCGCTTTCGCTTCTGGATGAACAAATTTATCAGC	1673
DB	1213	ATTGTCGCATCAGAAATTTGTTGAAACTTTTGTCTGGCTGGGAGATTTATTTGGA	1272
QY	1674	TATTCAAAGTCATTTCAGGACATCTTTTGGCAAGCATGCTTCAGAGAAACCTGACTCG	1733
DB	1273	TATTCAGATGGCTTCGCGGATTTTATGTCAAGTTGGTTTACAAAGAGGCCAGATGGT	1332
QY	1734	GAGATCCCTGTCATCAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCC	1793
DB	1333	GAATACTCTTCTCTAGTGGACAGACTGATCTAGTGAATTTCCAAATTTATGAAAGGTGCT	1392
QY	1794	CGTAAGATCCTTCATGAATTTGATCTTCAGCAGTTTATTTTATGGAGATGGAATACAAAAA	1853
DB	1393	AGGAGTGTATCATCAGTTGTTTCATGAGCATGCTGCTTATGGATTTGAAATACAGAAA	1452
QY	1854	CTCTTTGCTATGGAAATTTGTGAAGTATTAATAACCACTGCAGAGAAAGATATATCAGTGT	1913
DB	1453	CTATTTGCTGTTGCAATTTGCAAAAAATTAACAGCAGTTGTCAGAGAGATTTTATGGAGAT	1512
QY	1914	GATCATGACAGAGATATCTCTATAAATGCACTTTCAGTTTCAGATGTTTACGTTCCTACT	1973
DB	1513	GATCAGCAGCAGCAGTGTGCGTGACCTCTATCTGTCCAGTTCTTCACGCGACCTACT	1572
QY	1974	CTGGCTCGACATCTTATTTGAAGAGCAGATGTTATCTCTGTCAATTTACGAACTCTGCTA	2033
DB	1573	CTGGCTCGATGCTCTATCAGAGAAATACTTGATGAGCATTAATCATTTAGACTTTTATG	1632
QY	2034	GAAGTTTT-----ACCTGAGTACTTGGACAGGAAACAATAAATTTCAACTTCACGGGTTAT	2087
DB	1633	GATCAATTTGAGACATCCAGATGCCAGGCGAGATTTTCAGTTTGAACGATACATCTGTTA	1692
QY	2088	AGCCAGACAAATTTGGGAGAGTATATGCAATATATGTCACCTTAAGATATATCTGATC	2147
DB	1693	CAAGCTTCAAAATTTAGGAGAGTACAGAGCCTTATTTTAGATCTCAAGATGTTGTTAAT	1752
QY	2148	AGCAAAACCAATATGCAAGAAAGATTAAGAAATGCAAGTTCCTTTGAAGGTTTTTCGATCT	2207
DB	1753	AGCAAAACCACTGAATGTCAGATGAGCTGAGGCGAGATTTCTTAGAAGGTTTGTATGCC	1812
QY	2208	TTTTTGAAGATTTTACCTGTATGACAGGAAATGAAGAAATTCGAAGCAGGTTGGGCAA	2267
DB	1813	TTTTTGGAAATTTACTTAAATTTATGACAGGAATGATCCAAATTTACACGTCAGTAGGACAA	1872
QY	2268	CACATTTGAAGTGGATCTGATTTGGGCGCTGCCATTTGCTATACAGATGCAATTCAGAGAT	2327
DB	1873	CATATTTGAATTTGGAACAGAGTGGGAGCAGCCTTCACTACAAATTTGAATTTAACAT	1932
QY	2328	ATTTTACTCTGTTTCCAGAGTGTGTGCTGTGATGCAAGAACTCTTACTTGTGGCTTAT	2387
DB	1933	GTCAITTCATGATGACGAGCTGTTGCTTTCAGATGAAAGTGTTAATCGAAGCTTAC	1992
QY	2388	AAAGAAATGTCAAAAGCTGTGATGAGTGCAGTACCAAGTTTCATATCTAGTAGCAAGACA	2447
DB	1993	AAAGAAATGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2052
QY	2448	GT----AGTACATCGTGTGGACATAGTTTGGAAACAAAGTCTTACAGATATCTGAGAT	2504
DB	2053	ATCACCTAAGCAATTTTGGGACATTTTCAGTGGAACTATCAGATATCTGTTTCCCAAGAA	2112
QY	2505	CTTGTAGACATACATCTGCCACTCTCTAGGACCTTGTGCTGCTTCTCATGTACGTTTAAGC	2564
DB	2113	AAAGTAGCATTCACCTCCAGTTTCTCGCTTATCTTGCAGGTTTACATGATTTATTAAGC	2172
QY	2565	AGGCTGGGTGCTGTTTCAAGACTGCATGAATTTGTGTTTTTGGAGGACTTTTCAAGTAGAG	2624

Db 2173 AAAAGTGAAGTGGCATATAAAATTTCCAGAGCTCTACCTCTAAGTGAACCTTAGCCACCC 2232  
Qy 2625 GTACTAGTGAATATCTTTAGTGTCTGTGTGGTGGTCCAGGTTGTCTGAGATG 2684  
Db 2233 ATGTTGTAGATAACACCTCTTAGATGTCTTGTGTGTGGCCAAAGTACATCCCGAAATG 2292  
Qy 2685 TGGCCGAAGAAATGGAGTGTCTCTTATAGCCAGGTGTTTATTAACAGATGTTAAGTGC 2744  
Db 2293 TGGAGAGAAATGGGTCTCTCTAGTAAACAGATTTATCTACCAATAGTGAATGTC 2352  
Qy 2745 AGAGAGAAATGTATGATAAAGATATCATCATGCTTCAGATTTGGTGCATCTTTAATGGAT 2804  
Db 2353 AGACGTGAGATGTTTGGACAAGSAGTGTAGTAAATGCTTCAGACAGGTGTCTCCATGATGAT 2412  
Qy 2805 CCCAATAGTCTTGTGTACTGTCTAGAGGTATGAACCTTGGCCGAGGCTTTTAAACAG 2864  
Db 2413 CCAATATCATTTCTGATGATCATGTCTAGCCCTTTTGAACCTTTATCAGATTTTTCAGTACT 2472  
Qy 2865 ACCATATCTACAAA- - - - - GACCAGGATTTGATTAACAATAT 2903  
Db 2473 CCAGACTATGGAAGAAAGATTTAGTTCGAGATTAACCATAGGATGTTTTCAGCAGAAC 2532  
Qy 2904 AATACACTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTTGGGTGAGGTTATGTA 2963  
Db 2533 AATACTCTANTAGAGAAATGCTATACCTCATTAATGCTTTGTTGGAGAGATTTAGT 2592  
Qy 2964 CCTGAGTGGGAAATGTGCCAAAGAGAGGTCAATGAGAGAAATCATTCACCTGCTT 3023  
Db 2593 CCTGAGTGGCAGGTAAATGCTACAGATGAATCAAGCGAGAGATTTATCCATCAGTTG 2652  
Qy 3024 TGCATTGAACCCATGCCACAGTGCCTTGCAGAAATTTACCTGAGATGAAGAAATAT 3083  
Db 2653 AGTATCAAGCCCTATGGCTCATAGTGAATTTGTTAAAGTCTTTACCTGAAGTGAAGAACAG 2712  
Qy 3084 GAAACTGCTGTAGAGAAATGTATAAACAAAGTGGCCACATTTAAGAAACACAGGTGTATCA 3143  
Db 2713 GAGACTGSCATGAGAGTGTAAATCGAAGCAGTGTGCCATTTCAAGAACTGGATTAACA 2772  
Qy 3144 GGCCATGAGTTTATGAATAAAGATGAATCACTGAAGACTTCAATATGTTACTTTTAT 3203  
Db 2773 GGACGAGCATGTATGAATCTGAACCAAGATGTGCCAAAGAGTTCAACTTGTATTTCTAT 2832  
Qy 3204 CATTTACTCAAAACCCAGCATAGCAGGCTGAAATATGAGAGAAAGAGAGAAACAA 3263  
Db 2833 CACTTTTCAAGGGCGAGAACAGTCCAAAGCGCAGAGAAAGCGCAACGGAATTTGAAGACAA 2892  
Qy 3264 GAAACAAAGATGAAGCATTTGCCGCCACACACCTCTGAAATCTGCCCTGCTTTTCAGC 3323  
Db 2893 AATAGAGAGATACAGCACTCCACCTCGGTGTGGCTTCCATTTCTGCCCTCTGTTTGA 2952  
Qy 3324 AAAGTGATTAACCTTCTCAACTGTGTATATCATGATGTATCATTTCTCAGGACCGTATTTGAG 3383  
Db 2953 AGCTGGTTAAACATTTTTCAGTCAAGTGTATGTTGTGCATCATGGGAACAAATCTGCAA 3012  
Qy 3384 CGGCAATAGACACAGATTTCACTTGTGGACCGAAGGATGCTCCAAATGGCTTTTCAAT 3443  
Db 3013 TGGGCTGTGGAAACATAATAGATATGCTTGTGTCAGAGTCCATGCTGCAAGGGTGTATCAT 3072  
Qy 3444 ATCTGGCATTTGGGTTTACTAGAGAGAGAGCAACAGCTTCAAAAGCTCTGTAAG- - - AA 3500  
Db 3073 TTAATTTGGCATGGCACTACAGAGAGAGAAACAAACATTTAGAGATGTCAACGGAAGCAT 3132  
Qy 3501 GAAGTAAATTTGACTTTTATCATAGGCTTCAAGATTTGGGAAGTTTCAGCCATGAAT- - - 3557  
Db 3133 GTAGTAAACATTTACTTCACTCAGAGATATCAAAACCTGTGTGAAGCCCAAAATTTCT 3192  
Qy 3558 - - - - - ATACAAATGCTTTTGGAAAACTCAAGAGAAATTTCCAGTTTGAAGGCCAGAG 3611  
Db 3193 CCTAGCATACTAGCTATGCTGGAAACACTACAAAATGCTCCCTACCTAGAAAGTCCACAA 3252  
Qy 3612 GACATGATAACGTGGATACTTCAGATGTTTGACACAGTGAAGCGATTTAAGAGAAATCT 3671  
Db 3253 GACATGATTCGGTGGATTTGAAGACTTTTAAATGCTGTTTAAAGATGAGGAGAGTTTCA 3312

Qy 3672 TGTTTAATTTGATGCAACCAATCAGGATCGGAATCTTATTAAGATGATGAGATTACTCAT 3731  
Db 3313 CCTACCAGTC- - - - - CGGTGGCAGAGACAGAAAGAACCAATAATGAAGAGAGTTCAAGG 3366  
Qy 3732 GATAAAGAAAAAGCAGAAAGAAAAAGCTGAAGCTGTAGGCTACATCGCCAGAG 3791  
Db 3367 GACAAAGCAAAAGCTGAGAGAGAGAAAGAGAGATTTCCAGACTGGCGAGAGAAAG 3426  
Qy 3792 ATCATGGCTCAGATGTCTGCCCTTACAGAAAACTTTCATTGAAACTCATATAACTGTAT 3851  
Db 3427 ATCATGGCTCAGATGTCTGAAATGTCAGCGCATTTTATTTGATGAAAAACAAGAACTCTTT 3486  
Qy 3852 GACATATCATCAGAAATGCTTGGGAAAGAGATTTCCATTTATGAGAGAGAGAGACCCCA 3911  
Db 3487 CAGCAGACATTTAGAACTGGATGCTCAACCTCTGCTGTTCTTGATCA- - - - - TAGCCCT 3540  
Qy 3912 GCAGTCAGTACTACTCTAGAAATTTGCTTTGGTCTCTAAACGGGTCCATCTGTTACTCAA 3971  
Db 3541 GTGGCTTCAGATATGACACTTACAGACTGGGTCCACACAACTCAGGTTCTCGAAACA 3600  
Qy 3972 AAGGAGTGTCTGACGTGTCATCTTTTGGCAAGAAACAGAGAGGTGAATAATAGAAAAATAT 4031  
Db 3601 AGACAAATTCGTTACATGTATATTTGTCAAGAGAGCAAGAGTTAAAGTGGAAAGCAGG 3660  
Qy 4032 GCCATGGTATTTATCGGCTGTGTCCAGAAATCTACTGCTTTAAACCCAGCA CAGGGGAAAA 4091  
Db 3661 GCAATGGTCTTTGGCAGCATTTGTTTTCAGAGATCAACTGTATTTATCAAAAAACAGAAATAA 3720  
Qy 4092 CCCATAGAACTCTCAGGAGAGCCCTAGACCCACTTTTATGAGTCCAGACTTCGCATAT 4151  
Db 3721 TTTATTCAGAACTCAGAAAAA- - - - - TATGATCCATTTATTCATGCACCTGATCTGCTTGT 3777  
Qy 4152 GGAATTTATACAGGAAGCTGTGGTCAATGATGACGAGGTGTCTGTCGACAGAGTATTTT 4211  
Db 3778 GGNACACACTAGTGTGCTGGGCACTTATGATGATGCCATTTGTTGGCAAGGTATTTT 3837  
Qy 4212 GAAGCTGTACAGCTGAGCTCTCAGAGCGCATTCATGTTGACCTTTT- - - - - T 4259  
Db 3838 GATTCGCTTCAAGCTAAGAAACAGCGAAGGCAACAGAGATTTACGCTTACATACGAGCTAT 3897  
Qy 4260 GACTTGGAAAGTGGAGATATCTTTGGCTCTTTGGCAATCTCTGTGCAATCTGTGATC 4319  
Db 3898 GATGTAGAAACGAGAGATTTCTTTGGCCCCCTTTGTGAATGCTTTGAGTAATCTGTTTAT 3957  
Qy 4320 CCCATTTATCTTTGCAACCTCAAAAGATAAACAGTGAAGATGCAGATGCTCTTGTCTCAA 4379  
Db 3958 CC- - - - - TCTGCTGCTTCTCCCAAGAAATATTTTAAACACAGTTAAATTTTTCAGAC 4011  
Qy 4380 CTTTGTACCTTGGCA CCGTGGATACAGACTGTTCTGGCCAGAAATATCAGGTTATAATA 4439  
Db 4012 CAACCAATCTGACTCAGTGGATTAAGAACATAATCTCAGCAAAATAAAGCAATTACAGTTT 4071  
Qy 4440 AGACATGCTTAAAGAGAAACCCAAATTCCTATTTTCTTATCAAGNATGGGAGATTC 4499  
Db 4072 CTTAGSAAAGAAAGTACTCTCTAATAAGTCCCTCTCAAAAGAAATTCAGAAATTTGGAT 4131  
Qy 4500 ACTTTGAGTTTCCATTCATCTCAGTTTGGCGTGTAGTCTTCGATTTAAATATTTCAAT 4559  
Db 4132 GAATTCAGCTCCCTGAAGGGTTCAGGCTGATTTTCTCTAAGATCCCTTATTTCTGAG 4191  
Qy 4560 AGCATCAAGGAAATGGTTATTTCTCTTTGCCCAACAATTTTATAGAAATTTGGAATTTGAAAGTG 4619  
Db 4192 AGCATAAAGAAATGCTAAACGACATTTTGGAACTGCTACCTACAAAGTGGGACTAAAGTT 4251  
Qy 4620 CCACCTGATGAAGGATCTCTGAGTCCCAATGCTGACCTGGAGACCTGGCTTTCACT 4679  
Db 4252 CATCCCAATGAAGAGTCTCTGTTTCCCAATAATGTTGGGGTAGCTGCGGTACACC 4311  
Qy 4680 ATCCAGGCAATTTGAAATCTTATTCGGAGATGAAGGAAACCTCTGTTTGGAGCACTTCAA 4739  
Db 4312 ATCCAAAGCATAGAAAGAAATTTTGAAGTATGAGATTAACCAATTTGTTTGGTCTTTTACCT 4371





QY 2247 ATCCGAAGACAGGTTGGGCAACACATTTGAAGTGGATCCTGATTTGGGAGGCTGCCATTGCT 2306  
Db 121 ATCCGAAGACAGGTTGGGCAACACATTTGAAGTGGATCCTGATTTGGGAGGCTGCCATTGCT 180  
QY 2307 ATACAGATCAATTTGAAGAATATTTTACTCATGTTTCCAAAGAGTGGTGTCTTTGTCATGAA 2366  
Db 181 ATACAGATCAATTTGAAGAATATTTTACTCATGTTTCCAAAGAGTGGTGTCTTTGTCATGAA 240  
QY 2367 GAACTCTTACTTGTGGCTTATAAAGAAATCTCAAAAGCTGTGATGAGGTGCAGTACCAAT 2426  
Db 241 GAACTCTTACTTGTGGCTTATAAAGAAATCTCAAAAGCTGTGATGAGGTGCAGTACCAAT 300  
QY 2427 TTCTATATCTAGTAGCAGACACTAGTACAACTGCTGGGCATAGTTTGGAAACAAAGTCC 2486  
Db 301 TTCTATATCTAGTAGCAGACACTAGTACAACTGCTGGGCATAGTTTGGAAACAAAGTCC 360  
QY 2487 TACAGAGTATCTGAGGATCTTGTAAAGCATATCTGCGCACTCTCTAGGACCTTGTGCTGT 2546  
Db 361 TACAGAGTATCTGAGGATCTTGTAAAGCATATCTGCGCACTCTCTAGGACCTTGTGCTGT 420  
QY 2547 TTCTATATCTAGTAGCAGGCTGGGTGCTGTTTCAAGACTGCATGAATTTGTGTCTTTT 2606  
Db 421 TTCTATATCTAGTAGCAGGCTGGGTGCTGTTTCAAGACTGCATGAATTTGTGTCTTTT 480  
QY 2607 GAGGACTTTCAAGTAGAGTACTAGTGGAAATCTCTTACGTTGTCTGGTGTGGTGGCC 2666  
Db 481 GAGGACTTTCAAGTAGAGTACTAGTGGAAATCTCTTACGTTGTCTGGTGTGGTGGCC 540  
QY 2667 CAGGTTGTGCTGAGATGTGGGGAAGAAATGACGTCTCTTATTAGCCAGGTGTTTTAT 2726  
Db 541 CAGGTTGTGCTGAGATGTGGGGAAGAAATGACGTCTCTTATTAGCCAGGTGTTTTAT 600  
QY 2727 TACCAAGATGTTAAGTGCAGAGAAGAAATGTAATGAATATCATCTGCTTCAGATT 2786  
Db 601 TACCAAGATGTTAAGTGCAGAGAAGAAATGTAATGAATATCATCTGCTTCAGATT 660  
QY 2787 GGTGCATCTTTAATGGATCCCAATTAAGTCTTGTGTTACTGTTACTTTCAGAGGTATGAATT 2846  
Db 661 GGTGCATCTTTAATGGATCCCAATTAAGTCTTGTGTTACTGTTACTTTCAGAGGTATGAATT 720  
QY 2847 GCCGAGGCTTTTAAACAAGACCATATCTACAAAGACCAAGGATTTGATTAACAATATAAT 2906  
Db 721 GCCGAGGCTTTTAAACAAGACCATATCTACAAAGACCAAGGATTTGATTAACAATATAAT 780  
QY 2907 ACATTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTGTTGGGTGAGGTTATGTACCT 2966  
Db 781 ACATTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTGTTGGGTGAGGTTATGTACCT 840  
QY 2967 GGAGTGGGAAATGTGACCAAGAAAGAGGTCAATGAGAGAAATCAATTCATTGCTTTGC 3026  
Db 841 GGAGTGGGAAATGTGACCAAGAAAGAGGTCAATGAGAGAAATCAATTCATTGCTTTGC 900  
QY 3027 ATTGAACCCATGCCACAGTGCCTATGCCAAATTTTACCTGGAATGAAATTAATGAA 3086  
Db 901 ATTGAACCCATGCCACAGTGCCTATGCCAAATTTTACCTGGAATGAAATTAATGAA 960  
QY 3087 ACTGCTTTAGAGAAATGTATAAACAAAGTGGCCACATTTAAGAAACCAAGGTGTATCAGGC 3146  
Db 961 ACTGCTTTAGAGAAATGTATAAACAAAGTGGCCACATTTAAGAAACCAAGGTGTATCAGGC 1020  
QY 3147 CATGAGTTTATGAATCTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 3206  
Db 1021 CATGAGTTTATGAATCTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 1080  
QY 3207 TACTTCAAAACCCAGCATAGCAGGCTGAAATATGTCAGAGAAAGAGGAAACAAAGAA 3266  
Db 1081 TACTTCAAAACCCAGCATAGCAGGCTGAAATATGTCAGAGAAAGAGGAAACAAAGAA 1140  
QY 3267 AACAAAGATGAAGCATTTGCCGCCACACCACTCTGTAATTTCTGCCCTGCTTTTCAGCAA 3326  
Db 1141 AACAAAGATGAAGCATTTGCCGCCACACCACTCTGTAATTTCTGCCCTGCTTTTCAGCAA 1200  
QY 3327 GTGATTAACTTCTCAACTGTGTATATCATGTATGATATCTCTCAGGACCGTATTTGAGCGG 3386

Db 1201 GTGATTAACTTCTCAACTGTGTATATCATGTATGATATCTCAGGACCGTATTTGAGCGG 1260  
QY 3387 GCAATAGACACAGATTCTTAACCTTGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATT 3446  
Db 1261 GCAATAGACACAGATTCTCCACTTGTGGACCGAAGGGATGCTCCAAATGGATTGGGAGTT 1320  
QY 3447 CTGGCAT 3453  
Db 1321 CAGCCAT 1327

Search completed: July 8, 2005, 19:42:33  
Job time : 2927 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	6308	100.0	6308	4	US-09-724-126A-1	Sequence 1, Appli
2	5136	81.4	5205	4	US-09-724-126A-18	Sequence 18, Appli
3	4435.4	70.3	6395	4	US-08-982-956-1	Sequence 1, Appli
4	4435.4	70.3	6395	3	US-09-228-317-1	Sequence 3, Appli
5	1381.8	21.9	6300	4	US-09-724-126A-3	Sequence 3, Appli
6	1233.6	19.6	6089	4	US-09-724-126A-5	Sequence 5, Appli
7	996.2	15.8	1001	2	US-08-982-956-2	Sequence 2, Appli
8	996.2	15.8	1001	3	US-09-228-317-2	Sequence 2, Appli
9	103.4	1.8	670	4	US-09-370-767-174761	Sequence 14761, A
10	105.6	1.7	366	4	US-09-513-9990-936	Sequence 936, App
11	105.6	1.7	456	4	US-09-513-9990-938	Sequence 938, App
12	103.2	1.6	605	4	US-09-513-9990-937	Sequence 937, App
13	100.8	1.6	602	4	US-09-513-9990-14808	Sequence 14808, A
14	100.8	1.6	692	4	US-09-513-9990-14809	Sequence 14809, A
15	100.8	1.6	842	4	US-09-513-9990-14947	Sequence 14947, A
16	85.2	1.4	2182	4	US-09-370-767-13447	Sequence 13447, A
17	57.4	0.9	7218	1	US-08-232-463-14	Sequence 14, Appl
18	52.2	0.8	7218	1	US-08-232-463-14	Sequence 14, Appl
19	50.4	0.8	601	4	US-09-949-016-179951	Sequence 179951, A
20	50.4	0.8	101356	4	US-09-949-016-12364	Sequence 12364, A
21	50.4	0.8	101357	4	US-09-949-016-16924	Sequence 16924, A
22	49.2	0.8	932	4	US-09-573-080A-76	Sequence 76, Appl
23	49	0.8	1141	4	US-09-806-708B-22	Sequence 22, Appl
24	47.4	0.8	601	4	US-09-949-016-86450	Sequence 86450, A
25	47	0.7	129899	4	US-09-949-016-14684	Sequence 14684, A
26	47	0.7	148156	4	US-09-949-016-11776	Sequence 11776, A
27	46.2	0.7	95648	4	US-09-949-016-13139	Sequence 13139, A

Db 301 CACCCTGGAAACCACTTCTACTTTCTGTGTATGAATTTGACTACTCTAGCTGGATC 360  
Qy 361 CCGAGCTTTTGTATACATGTGCAAGTGGCCACGGGTAGAACTCTTAATAATAGAAT 420  
Db 361 CCGAGCTTTTGTATACATGTGCAAGTGGCCACGGGTAGAACTCTTAATAATAGAAT 420  
Qy 421 GTATGCAACAGTTCCTCAGCACCACCAACCCAGATATACAAACCAATTCAGCTACCAAGAGCTAC 480  
Db 421 GTATGCAACAGTTCCTCAGCACCACCAACCCAGATATACAAACCAATTCAGCTACCAAGAGCTAC 480  
Qy 481 GCCTGATAAATTTAGAGGGGAAAAAATCTCCAGTCCCTTCACGTCGTGACGCTTGC 540  
Db 481 GCCTGATAAATTTAGAGGGGAAAAAATCTCCAGTCCCTTCACGTCGTGACGCTTGC 540  
Qy 541 TCCGGGAAGCGGCGCGGAAGCACACTCCTCGAGTCTGCGTCAAAACCGGACTTCAGGGGCG 600  
Db 541 TCCGGGAAGCGGCGCGGAAGCACACTCCTCGAGTCTGCGTCAAAACCGGACTTCAGGGGCG 600  
Qy 601 TCGTAAAAAGTGTGCTCCCTGTCTCTCCGACCGGCCACAGGTTTCGCTTCGCTCTGGCGG 660  
Db 601 TCGTAAAAAGTGTGCTCCCTGTCTCTCCGACCGGCCACAGGTTTCGCTTCGCTCTGGCGG 660  
Qy 661 GGGGTCCGCAACTGCAAGCGTCAAGTTCCTCAAGATGGCGGACGAGGAGGCTGGAGGTA 720  
Db 661 GGGGTCCGCAACTGCAAGCGTCAAGTTCCTCAAGATGGCGGACGAGGAGGCTGGAGGTA 720  
Qy 721 CTGAGAGATGGAATCAGCGCGGAGTTACCCAGACCCCTCAGCGGCTGGGATCTTGGT 780  
Db 721 CTGAGAGATGGAATCAGCGCGGAGTTACCCAGACCCCTCAGCGGCTGGGATCTTGGT 780  
Qy 781 GGGATCAGCAAGTTCATTTTATCTCTCTCCGACCGGCCACAGGTTTCGCTTCGCTCTGGCGG 840  
Db 781 GGGATCAGCAAGTTCATTTTATCTCTCTCTCCGACCGGCCACAGGTTTCGCTTCGCTCTGGCGG 840  
Qy 841 AAATTTACTTGTCTGAAATGGACCCAGCTTGGAAAGCAGGAGAAAGTGTACAAATGT 900  
Db 841 AAATTTACTTGTCTGAAATGGACCCAGCTTGGAAAGCAGGAGAAAGTGTACAAATGT 900  
Qy 901 CAATATTCCTCACTGGAAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960  
Db 901 CAATATTCCTCACTGGAAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960  
Qy 961 AATTGAAGCAGATGCGGACATTTACGCTTTGTGGAGGGGTTTTCAAAAGTGGAGAGACA 1020  
Db 961 AATTGAAGCAGATGCGGACATTTACGCTTTGTGGAGGGGTTTTCAAAAGTGGAGAGACA 1020  
Qy 1021 CCTATTTCTGCAAGGATTTGCAATTTGATCCAAATGTGTACTCTGTATGGACTGCTTCC 1080  
Db 1021 CCTATTTCTGCAAGGATTTGCAATTTGATCCAAATGTGTACTCTGTATGGACTGCTTCC 1080  
Qy 1081 AGGACAGTGTTCATAAAAATCATCGTTACAGATGCATCTCTACTCTGATGGAGGGTTCT 1140  
Db 1081 AGGACAGTGTTCATAAAAATCATCGTTACAGATGCATCTCTACTCTGATGGAGGGTTCT 1140  
Qy 1141 GTGACTGTGGAGACACAGGCGATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200  
Db 1141 GTGACTGTGGAGACACAGGCGATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200  
Qy 1201 GAAAGCAGGTTACTATAAAGAGAAATTCACGCTGCTCCGTTGAATGAAGAGGTAATTTGCC 1260  
Db 1201 GAAAGCAGGTTACTATAAAGAGAAATTCACGCTGCTCCGTTGAATGAAGAGGTAATTTGCC 1260  
Qy 1261 AAGCCAGGAAATATTTCTCAGTGATAAATAATGTGCGTGAATAATGATATATGGGAAG 1320  
Db 1261 AAGCCAGGAAATATTTCTCAGTGATAAATAATGTGCGTGAATAATGATATATGGGAAG 1320  
Qy 1321 AGGAAAAAGAACTGCTCCTGAACTCCAGATAAGGGAGAAAAATCAAGAGATACTATTGT 1380  
Db 1321 AGGAAAAAGAACTGCTCCTGAACTCCAGATAAGGGAGAAAAATCAAGAGATACTATTGT 1380  
Qy 1381 TCCTTTTCAATGATGAACCACTTCAATATGACCGTCAATATACAGCCCTACAAAGAGCTC 1440  
Db 1381 TCCTTTTCAATGATGAACCACTTCAATATGACCGTCAATATACAGCCCTACAAAGAGCTC 1440

Qy 1441 TTGACTGTGAGCTCGCAGAGSCCAGTTTGCATACCACTGCCATTTGACAAAGAGGGTCTC 1500  
Db 1441 TTGACTGTGAGCTCGCAGAGSCCAGTTTGCATACCACTGCCATTTGACAAAGAGGGTCTC 1500  
Qy 1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGAAAGATATAAAGAGTC 1560  
Db 1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGAAAGATATAAAGAGTC 1560  
Qy 1561 ATTGAGAAATGTCTCTCAACATCCACTTCATGTAGAGTATTACACTCAGAGATTATGG 1620  
Db 1561 ATTGAGAAATGTCTCTCAACATCCACTTCATGTAGAGTATTACACTCAGAGATTATGG 1620  
Qy 1621 CTCAATCAGAAATTTGCTTTTGGCTCTTGGTCTCTGATGAAACAAATTTATGAGCTATTCAA 1680  
Db 1621 CTCAATCAGAAATTTGCTTTTGGCTCTTGGTCTCTGATGAAACAAATTTATGAGCTATTCAA 1680  
Qy 1681 GTGACTTTAGGCAGATCTTTTGCACAGCATGCCCTTAGAGAAGAACTGACTCGGAGATC 1740  
Db 1681 GTGACTTTAGGCAGATCTTTTGCACAGCATGCCCTTAGAGAAGAACTGACTCGGAGATC 1740  
Qy 1741 CCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAAGGTGCCGTAGA 1800  
Db 1741 CCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAAGGTGCCGTAGA 1800  
Qy 1801 TCCCTTCAATGAAATTCATCTTCAGCAGATTTTATATGGAGATGGAATACAAATACTCTTG 1860  
Db 1801 TCCCTTCAATGAAATTCATCTTCAGCAGATTTTATATGGAGATGGAATACAAATACTCTTG 1860  
Qy 1861 CTATGGAATTTGTGAAGTATTATAAACAACCTGCAAGAAATATATCAGTGATGATCATG 1920  
Db 1861 CTATGGAATTTGTGAAGTATTATAAACAACCTGCAAGAAATATATCAGTGATGATCATG 1920  
Qy 1921 ACAGAGTATCTCTATTAACCTGCACTTCAGTTGAGATGTTTACTGTTCTTACTCTGGCTC 1980  
Db 1921 ACAGAGTATCTCTATTAACCTGCACTTCAGTTGAGATGTTTACTGTTCTTACTCTGGCTC 1980  
Qy 1981 GACATCTTATTTGAAGAGCAGAAATGTTATCTGTCTATTACTGAACTCTCTAGAGTTT 2040  
Db 1981 GACATCTTATTTGAAGAGCAGAAATGTTATCTGTCTATTACTGAACTCTCTAGAGTTT 2040  
Qy 2041 TACCTGAGTACTCTGGACAGGAACAATAAATTCACCTCCAGGGTTATAGCCAGGACAAAT 2100  
Db 2041 TACCTGAGTACTCTGGACAGGAACAATAAATTCACCTCCAGGGTTATAGCCAGGACAAAT 2100  
Qy 2101 TGGGAAGATATATGCGATTAATGTGACCTTAAAGTATATCTGATCAGCAAAACCCACA 2160  
Db 2101 TGGGAAGATATATGCGATTAATGTGACCTTAAAGTATATCTGATCAGCAAAACCCACA 2160  
Qy 2161 TATGACACAGAAAGATTAAGAAATGCAATGAGTTTCCGATCTTTTGAAGATTC 2220  
Db 2161 TATGACACAGAAAGATTAAGAAATGCAATGAGTTTCCGATCTTTTGAAGATTC 2220  
Qy 2221 TTACTGTTATGACGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTTGAAGTGG 2280  
Db 2221 TTACTGTTATGACGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTTGAAGTGG 2280  
Qy 2281 ATCTGATTTGGAGGCTGCCATTCCTATACAGATGCAATTCGAAGATTTTACTCATGT 2340  
Db 2281 ATCTGATTTGGAGGCTGCCATTCCTATACAGATGCAATTCGAAGATTTTACTCATGT 2340  
Qy 2341 TCCAAAGAGTGTGTGCTTGTGATGAAGAACTTACTTGTGCTTTATAAAGATGTCA 2400  
Db 2341 TCCAAAGAGTGTGTGCTTGTGATGAAGAACTTACTTGTGCTTTATAAAGATGTCA 2400  
Qy 2401 AAGCTGTGATGAGGTGCAAGTCCAGTTCATATCTAGTAGCAAGACAGTAGTACAATCGT 2460  
Db 2401 AAGCTGTGATGAGGTGCAAGTCCAGTTCATATCTAGTAGCAAGACAGTAGTACAATCGT 2460  
Qy 2461 GTGACATAGTTTGGAAACAAAGTCTTACAGATATCTGAGGATCTTGTAGGATACATC 2520  
Db 2461 GTGACATAGTTTGGAAACAAAGTCTTACAGATATCTGAGGATCTTGTAGGATACATC 2520



QY 2521 TGCACCTCTCTAGGACCCCTTGTGCTCTTCACTAGTACGTTTAAAGCAGGCTGGGTGCTGTTT 2580  
Db 2521 TGCACCTCTCTAGGACCCCTTGTGCTCTTCACTAGTACGTTTAAAGCAGGCTGGGTGCTGTTT 2580  
QY 2581 CAAGACTGCATGAATTTGTGCTCTTGTAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640  
Db 2581 CAAGACTGCATGAATTTGTGCTCTTGTAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640  
QY 2641 CTTTACGTTGCTGTGTTGTTGCCAGGTTTGTCTGAGATGTGGCGAAGAAATGGAC 2700  
Db 2641 CTTTACGTTGCTGTGTTGTTGCCAGGTTTGTCTGAGATGTGGCGAAGAAATGGAC 2700  
QY 2701 TGTCTCTTATAGCCAGGTGTTTATTAACCAAGATGTTAAGTGCAGAGAAAGATGTATG 2760  
Db 2701 TGTCTCTTATAGCCAGGTGTTTATTAACCAAGATGTTAAGTGCAGAGAAAGATGTATG 2760  
QY 2761 ATAAAGATATCATCATGCTTCAGATTGTTGTCATCTTTAATGATCCCAATAGTCTTGT 2820  
Db 2761 ATAAAGATATCATCATGCTTCAGATTGTTGTCATCTTTAATGATCCCAATAGTCTTGT 2820  
QY 2821 TACTGGTACTTCAAGAGGTATGAATTTGCCAGGCTTTTAAACAAGACCATACTACAAAAG 2880  
Db 2821 TACTGGTACTTCAAGAGGTATGAATTTGCCAGGCTTTTAAACAAGACCATACTACAAAAG 2880  
QY 2881 ACCAGGATTTGATTAACCAATATACTAATAGAGAAATGCTTCAGGTCTCTCATCT 2940  
Db 2881 ACCAGGATTTGATTAACCAATATACTAATAGAGAAATGCTTCAGGTCTCTCATCT 2940  
QY 2941 ATATTGCGGTGAGGTTATGTACCTGGAGTGGGAAATGTACCAAGAGAGGTACAA 3000  
Db 2941 ATATTGCGGTGAGGTTATGTACCTGGAGTGGGAAATGTACCAAGAGAGGTACAA 3000  
QY 3001 TGAGAGAAATCAATTCATCTGCTTTGTCATTTGCAATGCCACACAGTGCCTCACTGA 3060  
Db 3001 TGAGAGAAATCAATTCATCTGCTTTGTCATTTGCAATGCCACACAGTGCCTCACTGA 3060  
QY 3061 ATTTACCTGAGAAATGAATTAATGAATCTGGCTTAGAGAAATGTATAAACAAAGTGGCCA 3120  
Db 3061 ATTTACCTGAGAAATGAATTAATGAATCTGGCTTAGAGAAATGTATAAACAAAGTGGCCA 3120  
QY 3121 CATTTAAGAAACCCAGGTGATCAGGCCATGGAGTTTATGAATGAATGAATCACTGA 3180  
Db 3121 CATTTAAGAAACCCAGGTGATCAGGCCATGGAGTTTATGAATGAATGAATCACTGA 3180  
QY 3181 AAGACTTCAATATGTACTTTTATCAATTTACTCCAAACCCAGCATAGCAGGCTGAACATA 3240  
Db 3181 AAGACTTCAATATGTACTTTTATCAATTTACTCCAAACCCAGCATAGCAGGCTGAACATA 3240  
QY 3241 TGCAGAGAAAGGAGAAACCAAGAAACAAAGATGAAGCATTTGCCGCCACCAACCTC 3300  
Db 3241 TGCAGAGAAAGGAGAAACCAAGAAACAAAGATGAAGCATTTGCCGCCACCAACCTC 3300  
QY 3301 CTGAATTCGCCCTCTTTCAGCAAGATGATTAACCTTCTCACTGTGATATCATGATGT 3360  
Db 3301 CTGAATTCGCCCTCTTTCAGCAAGATGATTAACCTTCTCACTGTGATATCATGATGT 3360  
QY 3361 ACATTCTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCACTTTGTGGACCGAAG 3420  
Db 3361 ACATTCTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCACTTTGTGGACCGAAG 3420  
QY 3421 GGATGCTCAAAATGCTTTTCAATATCTGGCATTTGGTTTATAGAGAGAGCAACAGC 3480  
Db 3421 GGATGCTCAAAATGCTTTTCAATATCTGGCATTTGGTTTATAGAGAGAGCAACAGC 3480  
QY 3481 TTCAAAAGCTCTGAGAGAGATTAACATTTGACTTTTATCATAGGCTTCAAGATTTG 3540  
Db 3481 TTCAAAAGCTCTGAGAGAGATTAACATTTGACTTTTATCATAGGCTTCAAGATTTG 3540  
QY 3541 GAAGTTACGACCATGAATATACAAATGCTTTTGGAAATCTCAAGGAAATTTCCCGATTAG 3600  
Db 3541 GAAGTTACGACCATGAATATACAAATGCTTTTGGAAATCTCAAGGAAATTTCCCGATTAG 3600  
QY 3601 AAGGCCAAGAGACATGATAACGTTGGATACCTTTCAGATGTTTGAACACAGTGAAGCATTA 3660

Db 3601 AAGGCCAAGAGACATGATAACGTTGGATACCTTTCAGATGTTTGAACACAGTGAAGCATTA 3660  
QY 3661 GAGAAAAATCTTTTAAATTTAGTGTAGCAACCATCAGGATCGGAATCTATTAAAGATGATG 3720  
Db 3661 GAGAAAAATCTTTTAAATTTAGTGTAGCAACCATCAGGATCGGAATCTATTAAAGATGATG 3720  
QY 3721 AGATTACTCATGATAAAGAAAGCAGAAACGAAAGAAAGAGCTGAAAGCTGTAGGCTAC 3780  
Db 3721 AGATTACTCATGATAAAGAAAGCAGAAACGAAAGAAAGAGCTGAAAGCTGTAGGCTAC 3780  
QY 3781 ATCCGCCAAGATCATATGCTCAGATGCTGTCTTACAGAAAACTTCACTTGAACCTCAT 3840  
Db 3781 ATCCGCCAAGATCATATGCTCAGATGCTGTCTTACAGAAAACTTCACTTGAACCTCAT 3840  
QY 3841 AACTCATGTATGACAAATACATCAGAAATGCTTGGGAAAGAGATTTCCATTTAGAGGAAG 3900  
Db 3841 AACTCATGTATGACAAATACATCAGAAATGCTTGGGAAAGAGATTTCCATTTAGAGGAAG 3900  
QY 3901 AGAGCAACCCAGCAGTCACTCTAGAAATGCTTTGGGCTCTTAAACGGGGTCCAT 3960  
Db 3901 AGAGCAACCCAGCAGTCACTCTAGAAATGCTTTGGGCTCTTAAACGGGGTCCAT 3960  
QY 3961 CTGTTACTGAAAAGAGGTGCTGAGTGCATCTTGGCCAAAGAGAAACAGGAGGTGAAA 4020  
Db 3961 CTGTTACTGAAAAGAGGTGCTGAGTGCATCTTGGCCAAAGAGAAACAGGAGGTGAAA 4020  
QY 4021 TAGAAAAATATGCTTATTTATCGGCTGTGTCCAGAAATCTACTGCCCTTAAACCCAGC 4080  
Db 4021 TAGAAAAATATGCTTATTTATCGGCTGTGTCCAGAAATCTACTGCCCTTAAACCCAGC 4080  
QY 4081 ACAGGGGAAAAACCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTTCATGATCCAG 4140  
Db 4081 ACAGGGGAAAAACCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTTCATGATCCAG 4140  
QY 4141 ACTTGGCATATGGAACCTTATACAGAGAGCTGTGTCATGTAATGCAACGAGTGTCTGC 4200  
Db 4141 ACTTGGCATATGGAACCTTATACAGAGAGCTGTGTCATGTAATGCAACGAGTGTCTGC 4200  
QY 4201 AGAAGTATTTTGAAGTGTACAGCTGTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260  
Db 4201 AGAAGTATTTTGAAGTGTACAGCTGTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260  
QY 4261 ACTTGGAAAGTGGAGAAATATCTTTGCCCTCTTTGCAATCTCTGTGCAATACTGTGATCC 4320  
Db 4261 ACTTGGAAAGTGGAGAAATATCTTTGCCCTCTTTGCAATCTCTGTGCAATACTGTGATCC 4320  
QY 4321 CCATTATTTCTTTCGAACCTCAAAAGATAAACAGTGAAGTGCAGATGCTCTTGTCTAAC 4380  
Db 4321 CCATTATTTCTTTCGAACCTCAAAAGATAAACAGTGAAGTGCAGATGCTCTTGTCTAAC 4380  
QY 4381 TTTTGAACCTGGCAGCGTGGATACAGCTGTTCTGGCCAGAAATATCAGGTTTATATATA 4440  
Db 4381 TTTTGAACCTGGCAGCGTGGATACAGCTGTTCTGGCCAGAAATATCAGGTTTATATATA 4440  
QY 4441 GACATGCTTAAAGGAGAAACCCCAATTTCTTCTTAAATCAAGGAATGGAGATTTCTA 4500  
Db 4441 GACATGCTTAAAGGAGAAACCCCAATTTCTTCTTAAATCAAGGAATGGAGATTTCTA 4500  
QY 4501 CTTTGGAGTTCCATTTCCATCTGAGTTTGGGCTGTGATCTTCGATTAATATTTCAATA 4560  
Db 4501 CTTTGGAGTTCCATTTCCATCTGAGTTTGGGCTGTGATCTTCGATTAATATTTCAATA 4560  
QY 4561 GCATCAAGGAATGTTTATCTCTTTGCAACAAATTTATAGAAATGGATTTGAAGTGC 4620  
Db 4561 GCATCAAGGAATGTTTATCTCTTTGCAACAAATTTATAGAAATGGATTTGAAGTGC 4620  
QY 4621 CACCTGATGAAAGGAGTCTCGAGTCCCACTGCTGAGCCTGGAGCACTCGCGCTTCACTA 4680  
Db 4621 CACCTGATGAAAGGAGTCTCGAGTCCCACTGCTGAGCCTGGAGCACTCGCGCTTCACTA 4680  
QY 4681 TCCAGGCAATTTGAAAATCTATTGGGAGATGAAGAAACCTCTGTGTTGGAGCATTTCAAA 4740

Db 4681 TCCAGGCAATTGAAAATCTATTGGAGATGAAGGAAAACCTCTGTTTGGAGCACTTCAA 4740  
Qy 4741 ATAGGACGATTAATGGTCTGAAGCAATTAATGCACTTTGCACTGTCACAGAGGATTACCT 4800  
Db 4741 ATAGGACGATTAATGGTCTGAAGCAATTAATGCACTTTGCACTGTCACAGAGGATTACCT 4800  
Qy 4801 GTCCCTCAGTCTCTGATACAGAAACATCTGGTTCGTCTCTATCAGTCTGTTCTTCTTAACA 4860  
Db 4801 GTCCCTCAGTCTCTGATACAGAAACATCTGGTTCGTCTCTATCAGTCTGTTCTTCTTAACA 4860  
Qy 4861 TAAATACAGAGATACACATGCCCTTCTGCTATAGATCTGTTCTCATGTTGTTGGGTG 4920  
Db 4861 TAAATACAGAGATACACATGCCCTTCTGCTATAGATCTGTTCTCATGTTGTTGGGTG 4920  
Qy 4921 CTGTTGATGATTCACATCTGTTATTTGGGATGACCCCTGTTGATCTGAGCCCTTCTTTCAG 4980  
Db 4921 CTGTTGATGATTCACATCTGTTATTTGGGATGACCCCTGTTGATCTGAGCCCTTCTTTCAG 4980  
Qy 4981 TTAGTCTCTTCTTATACACACCTTTATCTCTTCCATTTGATCACCATGGCACATGCTTTC 5040  
Db 4981 TTAGTCTCTTCTTATACACACCTTTATCTCTTCCATTTGATCACCATGGCACATGCTTTC 5040  
Qy 5041 AGATACTACTTACAGTAGACACAGGCCCTACCCCTTGTCTCAGGTTCAAGAAGACAGTGAAG 5100  
Db 5041 AGATACTACTTACAGTAGACACAGGCCCTACCCCTTGTCTCAGGTTCAAGAAGACAGTGAAG 5100  
Qy 5101 AGGCTCATTCGGCATCTCTTCTTTTGCAGAAATTTCTCAATATACAAAGTGGCTCCATTG 5160  
Db 5101 AGGCTCATTCGGCATCTCTTCTTTTGCAGAAATTTCTCAATATACAAAGTGGCTCCATTG 5160  
Qy 5161 GGTGGATATTCCTGGCTGGTATTTGTTGGTCTCACTGAAGAAATGGCATACCCCTTATC 5220  
Db 5161 GGTGGATATTCCTGGCTGGTATTTGTTGGTCTCACTGAAGAAATGGCATACCCCTTATC 5220  
Qy 5221 TTGCTGTGCTGCAATGTTTTTCCATTTTCACTATTTAATTGGGGTAACTCCGCTCAGGAACTGC 5280  
Db 5221 TTGCTGTGCTGCAATGTTTTTCCATTTTCACTATTTAATTGGGGTAACTCCGCTCAGGAACTGC 5280  
Qy 5281 ATACCAATTCGACAGAGAGTACAGTGCACTCTGTAGCTATCTATCTTACCTACAA 5340  
Db 5281 ATACCAATTCGACAGAGAGTACAGTGCACTCTGTAGCTATCTATCTTACCTACAA 5340  
Qy 5341 ATTTGTTCTGCTCTTCCAGGAATATTTGGGATCTGTAAGGCCCTTGTCTCCAGAGTGGT 5400  
Db 5341 ATTTGTTCTGCTCTTCCAGGAATATTTGGGATCTGTAAGGCCCTTGTCTCCAGAGTGGT 5400  
Qy 5401 GTGCAGATCTGCTCTTACTTAACTGTTTGAAGCAAAAAACACCGTGTGAGGTACCCCTA 5460  
Db 5401 GTGCAGATCTGCTCTTACTTAACTGTTTGAAGCAAAAAACACCGTGTGAGGTACCCCTA 5460  
Qy 5461 GAAAAAGAAATAGTTTGTATAGAGCTTCTGTAGTATAGTGGCTCTCTGATCAAGCTT 5520  
Db 5461 GAAAAAGAAATAGTTTGTATAGAGCTTCTGTAGTATAGTGGCTCTCTGATCAAGCTT 5520  
Qy 5521 CTCATTTCAGGTGCCAGGCTCTGACAGATGATGAGCAAGCATCTGTCCTCTGCTCTT 5580  
Db 5521 CTCATTTCAGGTGCCAGGCTCTGACAGATGATGAGCAAGCATCTGTCCTCTGCTCTT 5580  
Qy 5581 TCTGTGGGCTATATCTATGTTCTCAGAAATTTGTCGAGGAAATTTGTAACGGGGAAG 5640  
Db 5581 TCTGTGGGCTATATCTATGTTCTCAGAAATTTGTCGAGGAAATTTGTAACGGGGAAG 5640  
Qy 5641 AGGTTGAGCTTGCATTTTTCAGCACTTCACTGTGGAGCCGGAGTCTGCAATTTTCTTAA 5700  
Db 5641 AGGTTGAGCTTGCATTTTTCAGCACTTCACTGTGGAGCCGGAGTCTGCAATTTTCTTAA 5700  
Qy 5701 AAATCAGAGAAATGCGAGTGGTCTGTTGAAGGTAAGCCAGAGGCTGTGCTATCCAG 5760  
Db 5701 AAATCAGAGAAATGCGAGTGGTCTGTTGAAGGTAAGCCAGAGGCTGTGCTATCCAG 5760  
Qy 5761 CTCCTTACTTGGATGAATAGGAGAAACAGACCTTGGCTGAGAGGGGCAACCCCTTTC 5820  
Db 5761 CTCCTTACTTGGATGAATAGGAGAAACAGACCTTGGCTGAGAGGGGCAACCCCTTTC 5820

Qy 5821 ATTTATCTCGAGCGGTATCGGAAGCTCCATTTGGTCTGCGCAACAACACATGCAATTATAG 5880  
Db 5821 ATTTATCTCGAGCGGTATCGGAAGCTCCATTTGGTCTGCGCAACAACACATGCAATTATAG 5880  
Qy 5881 AAGAGATTTGCTAGAGCCAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTTAC 5940  
Db 5881 AAGAGATTTGCTAGAGCCAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTTAC 5940  
Qy 5941 TGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000  
Db 5941 TGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000  
Qy 6001 AAAATTTATGGAAGAACTTTCTGAGGCTGGGAAAGTATTGGAGGTTCTTTTCTCCATGTC 6060  
Db 6001 AAAATTTATGGAAGAACTTTCTGAGGCTGGGAAAGTATTGGAGGTTCTTTTCTCCATGTC 6060  
Qy 6061 CAGGTTCACTTACATCAATATAATTTCTTAAATGGAGTATTGCTTTCAATTAGCAACA 6120  
Db 6061 CAGGTTCACTTACATCAATATAATTTCTTAAATGGAGTATTGCTTTCAATTAGCAACA 6120  
Qy 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180  
Db 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180  
Qy 6181 ATTTATTTCCCTTCCATTAATTTGCTCTCATTTTCAATTTTATTTTCACTTGGTAGATGAAG 6240  
Db 6181 ATTTATTTCCCTTCCATTAATTTGCTCTCATTTTCAATTTTATTTTCACTTGGTAGATGAAG 6240  
Qy 6241 TCAGGTCACAGATTTGTAGACATTTATGTTTGGTTAACTCTTCTGCAATTTTGTATTT 6300  
Db 6241 TCAGGTCACAGATTTGTAGACATTTATGTTTGGTTAACTCTTCTGCAATTTTGTATTT 6300  
Qy 6301 GGTGTTTT 6308  
Db 6301 GGTGTTTT 6308

## RESULT 2

US-09-724-126A-18  
; Sequence 18, Application US/09724126A  
; Patent No. 6706505  
; GENERAL INFORMATION:  
; APPLICANT: Han, Hui-Quan  
; APPLICANT: Kwak, Keich  
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family  
; FILE REFERENCE: 01017/35966A  
; CURRENT APPLICATION NUMBER: US/09/724,126A  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,211  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 18  
; LENGTH: 5205  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (662)  
; OTHER INFORMATION: n = A or T or G or C  
; NAME/KEY: misc feature  
; LOCATION: (668)  
; OTHER INFORMATION: n = A or T or G or C  
US-09-724-126A-18

Query Match 81.4%; Score 5136; DB 4; Length 5205;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;  
Qy 696 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGGAGTTACCCAG 755  
Db 1 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGGAGTTACCCAG 60

QY 756 ACCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATATCTGCTTCTTG 815  
DB 61 ACCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATATCTGCTTCTTG 120  
QY 816 CATCATTTGGCAAAATTTGGTGGCAAGAAATTTACTTGTCTGAAATGAGCCAGACTTGGAA 875  
DB 121 CATCATTTGGCAAAATTTGGTGGCAAGAAATTTACTTGTCTGAAATGAGCCAGACTTGGAA 180  
QY 876 AAGCAGGAGAAAGTGTACAAATGCAATTAATCACTCCACTGGAAATGTTACTTATTTGGA 935  
DB 181 AAGCAGGAGAAAGTGTACAAATGCAATTAATCACTCCACTGGAAATGTTACTTATTTGGA 240  
QY 936 GAAGATCCAGATATTGCTTAGAGAAATTTGAAGCAGTGGAGCAATTCAGCTTTGTGGG 995  
DB 241 GAAGATCCAGATATTGCTTAGAGAAATTTGAAGCAGTGGAGCAATTCAGCTTTGTGGG 300  
QY 996 AGGGTTTTCAAAGTGGAGACAACCTATTCTTGCAGGGATTTGTCAATTCATCCAACA 1055  
DB 301 AGGGTTTTCAAAGTGGAGACAACCTATTCTTGCAGGGATTTGTCAATTCATCCAACA 360  
QY 1056 TGTGTACTCTGTATGAGCTGCTTCCAGGACAGTGTTCATAAAATCATGTTACAGATG 1115  
DB 361 TGTGTACTCTGTATGAGCTGCTTCCAGGACAGTGTTCATAAAATCATGTTACAGATG 420  
QY 1116 CATATCTTCTACTGGAGGGTCTGTGTACTGTGGAGACACAGAGGCAATGGAAAATCTGGC 1175  
DB 421 CATATCTTCTACTGGAGGGTCTGTGTACTGTGGAGACACAGAGGCAATGGAAAATCTGGC 480  
QY 1176 CCTTTTGTGTAATCATGAACCTCGAGAGCAGTACTATTAAGAGAAATTCAGCTGT 1235  
DB 481 CCTTTTGTGTAATCATGAACCTCGAGAGCAGTACTATTAAGAGAAATTCAGCTGT 540  
QY 1236 CGTGTGAATGAAGAGTAAATGTCCAGCCAGGAAATTAATTCCTTCACTGATGATAAATAT 1295  
DB 541 CGTGTGAATGAAGAGTAAATGTCCAGCCAGGAAATTAATTCCTTCACTGATGATAAATAT 600  
QY 1296 GTCTAGAAATGACTATATGGGAAGGAAAGAACTGCTTCCAGAACTCCAGATAAGG 1355  
DB 601 GTCTAGAAATGACTATATGGGAAGGAAAGAACTGCTTCCAGAACTCCAGATAAGG 660  
QY 1356 GAGAAAATGAAGATATTAATGTCTCTTTCAATGATGAACACCATTCATATGACCAAC 1415  
DB 661 KNR -----YYCNDHSDH 675  
QY 1416 GTCATATACAGCTCAAAAGAGCTCTTGACTGTGAGCTCGAGAGGCCAGTTGCATACC 1475  
DB 676 GTCATATACAGCTCAAAAGAGCTCTTGACTGTGAGCTCGAGAGGCCAGTTGCATACC 735  
QY 1476 ACTGCCATGACAAAGAGGGTCTGCGGCTGTATAAGCGGGAGCTTATGCTGCTTGCAG 1535  
DB 736 ACTGCCATGACAAAGAGGGTCTGCGGCTGTATAAGCGGGAGCTTATGCTGCTTGCAG 795  
QY 1536 GAAGCAAGGAGATATAAGAGTCAATTCAGAAATGCTCTCAACATCCACTTCATGTA 1595  
DB 796 GAAGCAAGGAGATATAAGAGTCAATTCAGAAATGCTCTCAACATCCACTTCATGTA 855  
QY 1596 GAAGTATTACACTCAGAGATATGCTCATCAGAAATTTGCTTTCGCTTGTGTTCCCTGG 1655  
DB 856 GAAGTATTACACTCAGAGATATGCTCATCAGAAATTTGCTTTCGCTTGTGTTCCCTGG 915  
QY 1656 ATGAACAAAATTTATGAGCTATTCAAGTGAATTTAGGAGAGTCTTTTGGCAAGCATGCCTT 1715  
DB 916 ATGAACAAAATTTATGAGCTATTCAAGTGAATTTAGGAGAGTCTTTTGGCAAGCATGCCTT 975  
QY 1716 AGAAGAAACCTGACTCGAGAAATCCCTGTCTCATAGAGGTTAATGCTTTGGATGCA 1775  
DB 976 AGAAGAAACCTGACTCGAGAAATCCCTGTCTCATAGAGGTTAATGCTTTGGATGCA 1835  
QY 1776 AAGCTTTATAAGGTCCTGAGATCTTCAATGAATTTGATCTTCAGCAGTCTTTTATG 1895  
DB 1036 AAGCTTTATAAGGTCCTGAGATCTTCAATGAATTTGATCTTCAGCAGTCTTTTATG 1955

QY 1836 GAGATGGAAATACAAAACCTCTTCTGCTATGGAATTTGTGAAGTATATATAACACTGCAG 1895  
DB 1096 GAGATGGAAATACAAAACCTCTTCTGCTATGGAATTTGTGAAGTATATATAACACTGCAG 1155  
QY 1896 AAAGAAATATATCAGTGATGATCATGACAGAACTATCTCTATAAATGCACTTTTCAAGTTCA 1955  
DB 1156 AAAGAAATATATCAGTGATGATCATGACAGAACTATCTCTATAAATGCACTTTTCAAGTTCA 1215  
QY 1956 ATGTTTACTGTTTCTTCTGCTGACATCTTATGAAGAGCAGAAATGTTATCTCTGTC 2015  
DB 1216 ATGTTTACTGTTTCTTCTGCTGACATCTTATGAAGAGCAGAAATGTTATCTCTGTC 1275  
QY 2016 ATTACTGAACTCTGCTAGAAAGTTTACTGAGTACTTTGGACAGAAATTAATAATCAAC 2075  
DB 1276 ATTACTGAACTCTGCTAGAAAGTTTACTGAGTACTTTGGACAGAAATTAATAATCAAC 1335  
QY 2076 TTCCAGGGTTATAGCAGACAAATTTGGAGAGTATATGCAAGTATATGTCAGCTTAAAG 2135  
DB 1336 TTCCAGGGTTATAGCAGACAAATTTGGAGAGTATATGCAAGTATATGTCAGCTTAAAG 1395  
QY 2136 TATATCTGATCAGCAAAACCCCAATATGGACAGAAAGATTAAGAAATGCACTTCTTGA 2195  
DB 1396 TATATCTGATCAGCAAAACCCCAATATGGACAGAAAGATTAAGAAATGCACTTCTTGA 1455  
QY 2196 GGTTCGATCTTTTGAAGATCTTACTCTGATGCGAGGAATGGAAATCCGAAGA 2255  
DB 1456 GGTTCGATCTTTTGAAGATCTTACTCTGATGCGAGGAATGGAAATCCGAAGA 1515  
QY 2256 CAGTTGGGCAACATTTGAAGTGGATCTGATTTGGAGGCTGCCATTCGTATACAGATG 2315  
DB 1516 CAGTTGGGCAACATTTGAAGTGGATCTGATTTGGAGGCTGCCATTCGTATACAGATG 1575  
QY 2316 CAATGAAGAAATATTTTACTCATGTTCCAAGAGTGGTGTGTTGATGAAGAACTCTTGA 2375  
DB 1576 CAATGAAGAAATATTTTACTCATGTTCCAAGAGTGGTGTGTTGATGAAGAACTCTTGA 1635  
QY 2376 CTTTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGGTGAGTACAGATTTTCAATCT 2435  
DB 1636 CTTTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGGTGAGTACAGATTTTCAATCT 1695  
QY 2436 AGTAGCAGACAGTAGTACAAATCGTGTGGAATAGTTTGGAAACAAAGCTCCTACAGATG 2495  
DB 1696 AGTAGCAGACAGTAGTACAAATCGTGTGGAATAGTTTGGAAACAAAGCTCCTACAGATG 1755  
QY 2496 TCTGAGGATCTTGTAAAGCATACATCTGCACTCTCTAGGACCTTGTGCTTCAATGTA 2555  
DB 1756 TCTGAGGATCTTGTAAAGCATACATCTGCACTCTCTAGGACCTTGTGCTTCAATGTA 1815  
QY 2556 CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGCAATGAATTTGTGCTTTTGGAGACTT 2615  
DB 1816 CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGCAATGAATTTGTGCTTTTGGAGACTT 1875  
QY 2616 CAAGTAGAGGTACTAGTGGAAATATCTTCAAGTGTGCTGTTGTTGGCCAGGCTGTT 2675  
DB 1876 CAAGTAGAGGTACTAGTGGAAATATCTTCAAGTGTGCTGTTGTTGGCCAGGCTGTT 1935  
QY 2676 GCTGAGATGTGGCAAGAAATGGAATGCTCTTATAGCCAGGTTGTTTATTAACCAAGAT 2735  
DB 1936 GCTGAGATGTGGCAAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1995  
QY 2736 GTTAAGTGCAGAGAAATGATATGATTAAGAAATCATCATGCTTCAAGATTTGGTGCATCT 2795  
DB 1996 GTTAAGTGCAGAGAAATGATATGATTAAGAAATCATCATGCTTCAAGATTTGGTGCATCT 2055  
QY 2796 TTAATGATCCCAATAGTCTTCTGTTACTGTAATTCAGAGGTATGAATTTGCCAGGCT 2855  
DB 2056 TTAATGATCCCAATAGTCTTCTGTTACTGTAATTCAGAGGTATGAATTTGCCAGGCT 2115  
QY 2856 TTTAAACAGACATATCTTACAAAAGACAGAGATTTGATTTAAACAAATATAATACACTAATA 2915  
DB 2116 TTTAAACAGACATATCTTACAAAAGACAGAGATTTGATTTAAACAAATATAATACACTAATA 2175  
QY 2916 GAAGAAATGCTTTCAGGCTCTCATCTATATTTGTTGGTGGAGCTTATGTTACCTGGAGTGGGA 2975

Db	2176		GAAGAAATGCTTCAAGGTCCTCATCTATATTTGTGGGTGAGCGTTATGTACTCTGGAGTGGGA	2235
Qy	2976	AATGTGACCAAGAAAGAGGTGCACAAATCAGAGAGAAATCAATTCATCTGCTTTTGCAATGGAACCC	3035	
Db	2236	AATGTGACCAAGAAAGAGGTGCANATGAGAGAAATCAATTCATCTGCTTTGCAATGGAACCC	2295	
Qy	3036	ATGCCACACAGTGCATTTGCCAAAAATTTTACCTGAGAAATGAAAATTAATGAAACTGGCTTA	3095	
Db	2296	ATGCCACACAGTGCATTTGCCAAAAATTTTACCTGAGAAATGAAAATTAATGAAACTGGCTTA	2355	
Qy	3096	GAGAAATGCTATAACAAAGTGGCCACATTTTAAGAAACACAGGTGTATCAGGCCATGGAGTT	3155	
Db	2356	GAGAAATGCTATAACAAAGTGGCCACATTTTAAGAAACACAGGTGTATCAGGCCATGGAGTT	2415	
Qy	3156	TATGAACTTAAAGATGAATCACTCAAGACCTTCAATATGTACTTTTATCAATTAATCTCCAAA	3215	
Db	2416	TATGAACTTAAAGATGAATCACTCAAGACCTTCAATATGTACTTTTATCAATTAATCTCCAAA	2475	
Qy	3216	ACCAGCATAGCAAGGCTGAAATATATGCAGAGAAAAAGAGAAAAACAAGAAACAAGAT	3275	
Db	2476	ACCAGCATAGCAAGGCTGAAATATATGCAGAGAAAAAGAGAAAAACAAGAAACAAGAT	2535	
Qy	3276	GAAGCATTTGCCGCCACCAACCACTCTCTGAATTTCTGCCCTGTTTCAGCAAGTGAATTAAC	3335	
Db	2536	GAAGCATTTGCCGCCACCAACCACTCTCTGAATTTCTGCCCTGTTTCAGCAAGTGAATTAAC	2595	
Qy	3336	CTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGGCAATAGAC	3395	
Db	2596	CTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGGCAATAGAC	2655	
Qy	3396	ACAGTTCTAACTTTGTGGACCGAAGGATGCTCCAAATGGCTTTTCAATATCTGGCATTG	3455	
Db	2656	ACAGTTCTAACTTTGTGGACCGAAGGATGCTCCAAATGGCTTTTCAATATCTGGCATTG	2715	
Qy	3456	GGTTTACTAGAAGAGAAACCAACAGCTTCAAAGAGCTCCTGAAGAGAAAGTAAACATTTGAC	3515	
Db	2716	GGTTTACTAGAAGAGAAACCAACAGCTTCAAAGAGCTCCTGAAGAGAAAGTAAACATTTGAC	2775	
Qy	3516	TTTTTATCATTAAGGCTTCAGATTTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAA	3575	
Db	2776	TTTTTATCATTAAGGCTTCAGATTTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAA	2835	
Qy	3576	AAACTCAAAGGAAATTCOCAGTTAGAAGGCCAGAAAGCAATGATAACGTCGATCTTCAG	3635	
Db	2836	AAACTCAAAGGAAATTCOCAGTTAGAAGGCCAGAAAGCAATGATAACGTCGATCTTCAG	2895	
Qy	3636	ATGTTTGACACAGTGAAGCGGANTTAAGAGAAAAATTTGTTTAAATGTPAGCAACCAATCA	3695	
Db	2896	ATGTTTGACACAGTGAAGCGGANTTAAGAGAAAAATTTGTTTAAATGTPAGCAACCAATCA	2955	
Qy	3696	GGATCGGAATCTATTAAAGATGATGAGATTTACTCATGATTAAGAAAAAGCAGAACGAAA	3755	
Db	2956	GGATCGGAATCTATTAAAGATGATGAGATTTACTCATGATTAAGAAAAAGCAGAACGAAA	3015	
Qy	3756	AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAAAGATCATGGCTCAGATGTCTGGCTTA	3815	
Db	3016	AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAAAGATCATGGCTCAGATGTCTGGCTTA	3075	
Qy	3816	CAGAAAAATCTTCAATTTGAAAATCTATAAATCTCATGTATGACAATAATCAGAAAAATGCTGGG	3875	
Db	3076	CAGAAAAATCTTCAATTTGAAAATCTATAAATCTCATGTATGACAATAATCAGAAAAATGCTGGG	3135	
Qy	3876	AAAGAGATTTCCATTATGAGGAAAGAGAGCAACCCAGCAGTTCAGTGAATTTCTAGAAAT	3935	
Db	3136	AAAGAGATTTCCATTATGAGGAAAGAGAGCAACCCAGCAGTTCAGTGAATTTCTAGAAAT	3195	
Qy	3936	GCTTTGGGTCCATAACCGGGTCCATCTGTTTACTGAAAGAGGCTGTCAGTGCATCCTT	3995	
Db	3196	GCTTTGGGTCCATAACCGGGTCCATCTGTTTACTGAAAGAGGCTGTCAGTGCATCCTT	3255	
Qy	3996	TGCCAAGAGAAACAGGAGGTGAAAAATAGAAAAATATGCCATGGTATTATCGGCTGTGTC	4055	





1189 ATCATGAACCTGGAGACAGGTACTATAAAGAGAAATTCACGCTGTCGCTTGAATGAAG 1248  
Db  
608 ATCCAGGCTTGGAGACAGGTACTACAAAGAGAGCTTACATTTGCCCATTTGAATGAAG 667  
Qy  
1249 AGGTAATTTTCCAAAGCCAGGAAAATATTTCTTCTCAGTGATATAAATATATCTGTAGAAATGA 1308  
Db  
668 AGGTGATTTGCTCAAGCCAGGAGAAATATTTCCCTTCCGTTGATATAAATACATTTGTAGAAATGA 727  
Qy  
1309 CTATATGGAGAGAGAAAAGAAAGAACTGCTCTGCTGAACTCCAGATAGGAGAGAAAATGAAG 1368  
Db  
728 CTATATGGAGAGAGAAAAGAAATTTGCTCTGCTGAACTCCAGATAGGAGAGAAAATGAAG 787  
Qy  
1369 GATACATTTTGTGCTCTTCAATGATGAACACCAATTCATATAGCACACCTCATATACAGCC 1428  
Db  
788 GATACATTTTGTGCTCTTCAACGNATGAGACCAATTCGTATGATCATGTGATCTACAGTC 847  
Qy  
1429 TACAAGAGCTTTGGAATGTGAGCTGCGAGAGGCCAGTTGATACCACTGCCATTTGACA 1488  
Db  
848 TGCAGAGAGCTCTAGATTTGCGAGCTTGCAGAGGCCACAGCTGCACACGACTGCCATCGACA 907  
Qy  
1489 AAGAGGCTCGCTGGGCTGTTAAAGCGGAGCTTATGCTGCTTGCCTGCGAGGAACCAAGGAG 1548  
Db  
908 AAGAGGCTCGCGGCTGCTCAAGCAGGTGTATGCGCACTTGCAGGAAGCAAGGAGG 967  
Qy  
1549 ATATAAGAGTCATTCAGAAAATGTCTCTCAACATCCCACTTCATGTAGAGATTAATCACT 1608  
Db  
968 ATATAAGAGTCATTCAGAGAACGTCTCTCAGACCCCTCCATCCATGTGAAATGTGCACT 1027  
Qy  
1609 CAGAGATTTATGGCTCATCAGAAAATTTGCTTTGCGTCTTTGGTCTTCTGGATGAACAAAATTA 1668  
Db  
1028 CCGTGGTTATGGCTCACAGAAAATTCGCTCTGCGCTTTGGCTCTCGATGAACAAAATTA 1087  
Qy  
1669 TGAGCTATTCAGTCACTTTAGGCAGATCTTTTGGCAAGCATGCTTAGAGAGAACCTG 1728  
Db  
1088 TGAGCTATTCAGTCACTTTAGACAGATATTTTGGCAGGCTTGCCTCGTAGAGAACCTG 1147  
Qy  
1729 ACTGGAGAACTCCCTGCTCATAGCAGGTAAATGCTTTGGGATGCAAGCTTTATAAAG 1788  
Db  
1148 GCTCTGAANAATCCCTGCTTTATAGCAGACTAATGCTTTGGGATGCAANAATTTATAAAG 1207  
Qy  
1789 GTGCCGCTPAAGATCCTTCATGAATTTGATCTTTAGCAGATTTTTTTATGAGATGGAATACA 1848  
Db  
1208 GTGCCGCTPAAGATCCTTCATGAATTTGATCTTTAGTATTTTTTTATGAGATGGAATACA 1267  
Qy  
1849 AAAAATCTTTGCTATGGAATTTGTGAAGTATTTATAAACAATGCAAGAGATATATCA 1908  
Db  
1268 AAAAATCTTTGCTATGGAATTTGTGAAGTATTTATAAACAATGCAAGAGATATATCA 1327  
Qy  
1909 GTGATGATCATGACAGAAAGTATCTCTATAAATGCACTTTTCAAGTTTCAAGATTTTACTGTTT 1968  
Db  
1328 GCGACGACACAGAGAGAGCATCTCCATTAACGCTGCTGCTGCGAGATGCTCACCGTCC 1387  
Qy  
1969 CTACTCTGGCTCGACATCTTATGGAAGCAGAAATTTTATCTGTCTATTACTGAAATC 2028  
Db  
1388 CGACCTTTGCGCCGCGCATCTTATGGAAGCAGAAATTTTATCTGTCTATTACTGAAAGC 1447  
Qy  
2029 TGCTAGAAATTTTACTGAGTACTTTGGACAGGAACAATAAATTCAACTTCCAGGGTTATA 2088  
Db  
1448 TGCTAGAAATTTTACTGAGTACTTTGGACAGGAACAATAAATTCAACTTCCAGGGTTATA 1507  
Qy  
2089 GCCAGACAAAATTTGGGAAGAGTATATGAGTAAATGAGCTTAAGTATATCTGTATCA 2148  
Db  
1508 GCCAGACAAAATTTGGGAAGAGTCTTAGCGAGTATATGATGACCTTAAGTATATCTGTATTA 1567  
Qy  
2149 GCAAAACCAATATGGAAGAGATTAAGAAATGCAATTTCTTGAAGTTTTCGATCTT 2208  
Db  
1568 GCAAGCTGCTCATATGGAAGAGATTAAGAAATGCAATTTCTTGAAGTTTTCGATCTT 1627  
Qy  
2209 TTTTGAAGATTTTACTGATGATGAGGAAATGGAAGAAATCGAGAGACAGTTTGGGCAAC 2268  
Db  
1628 TTTCTGAAGATTTTACTGATGATGAGGAAATGGAAGAAATCGAGAGACAGTTTGGGCAAC 1687  
Qy  
2269 ACAATGAATGATCTCTGATTTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAATA 2328

1688 ACATTTGAAGTGGACCCCTGACTGGGAGGCTGCCATTCAGATGCAACTAAGAATA 1747  
Qy  
2329 TTTTACTCATGTTTCCAGAGCTGGTGTGCTTGTGATGAGAACTCTTACTTTGTGCTTATA 2388  
Db  
1748 TTTTGTCTCATGTTTCCAAAGAGTGGTGTGCTTGTGATGAAGATCTCTTACTTGTGCTTATA 1807  
Qy  
2389 AAGAAATGTCACAAAGCTGTGATGAGGTGCAATCCAGTTTTCATATCTAGTAGCAAGACAG 2448  
Db  
1808 AAGAAATGTCACAAAGCTGTGATGAGGTGCAATCCAGTTTTCATGTCAGTACCAAGACAG 1867  
Qy  
2449 TAGTACAATCGTGTGGAATAGTTTGGAAAACAAAGTCTCAGAGTATCTGAGGATCTTG 2508  
Db  
1868 TAGTGCAATTTGTCGGTTCATAGTCTCGAAAACCAATCTCAAAAGTGTCTGAGGACCTTG 1927  
Qy  
2509 TAAGCATACATCTGCCACTCTTAGGACCTTGTGCTGCTTCAATGATAGTTTAAAGCAGC 2568  
Db  
1928 TAAGCATACACCTGCCACTCTTAGAACAATGCTGCTGCTTCAATGATAGTTTAAAGCAGC 1987  
Qy  
2569 TGGGTGCTGTTTCAAGACTGCAATGAAATTTGTGTCTTTTGAGGACTTTTCAAGTAGAGGTAC 2628  
Db  
1988 TAGTGTCTATTTCAAGACTGCAATGAAATTTGTGCTTTTGTGACAGCTTTTCAAGTAGAGGTCC 2047  
Qy  
2629 TAGTGAATATCTCTTACGTTGCTGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2688  
Db  
2048 TGGTGGAGTACCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2107  
Qy  
2689 GAAGAAATGCACTGCTCTTATTAGCCAGGTGTTTATTATCCAGAGATGTTTAAAGTGCAGAG 2748  
Db  
2108 GAAGAAATGCACTGCTCTCTCATCAGCAGGTGTTTCTATTATCAAGATGTTTAAATGTCAGGG 2167  
Qy  
2749 AAGAAATGTTATGATAAAGATATCATCATGCTTCAGATTTGGTGCATCTTTAATGATGCCA 2808  
Db  
2168 AGGAAATGTTACGATAAAGATATCATCATGCTTCAGATTTGGAGCATCTTATGATGCCA 2227  
Qy  
2809 ATAAAGTTCTTGTACTGCTACTTCAGAGGTATGAATTCGCGAGGCTTTTAAACAAGACCA 2868  
Db  
2228 ACAAGTTCTTGTACTGCTACTTCAGAGATATGAATTCATGATGCTTTTAAACAAGACCA 2287  
Qy  
2869 TATCTACAAAAGACCAAGATTTGATTAACAATATATATCACTTAATAGAGAAATGCTTC 2928  
Db  
2288 TATCCAAAAGACCAAGATTTGATTAACAAGATATATCACTTAATAGAGAAATGCTTC 2347  
Qy  
2929 AGGTCTCTCATCTATTTTGTGGTGCAGCTTATGATCTGAGGTGGGAAATGTCACCAAG 2988  
Db  
2348 AGGTCTCTCATCTATTTTGTGGAGAACTTATGATGCTGCTGAGTGGGAAATGTTTACCAG 2407  
Qy  
2989 AAGAGGTCCAAATGAGAGAAATCATCTTCTGCTTTGCAATTTGAAACCCATGCCACACAGTG 3048  
Db  
2408 AGGAGGTATTAATGAGAGAGATTAATCTCACTTCTTGTGATTTGAGCCCATGCCACACAGTG 2467  
Qy  
3049 CCATTTGCCAAAATTTTACCTGAGAAATGAAATATATGAAATCTGGCTTAGAGAAATGTCATAA 3108  
Db  
2468 CCATTTGCCAAAATTTTACCTGAGAAATGAAATATATGAAATCTGGCTTAGAGAAATGTCATAA 2527  
Qy  
3109 ACAAAGTGGCCACATTTTAAAGAAACAGGTGATCAGGCCATGAGGATTTTATGAACTTAAAG 3168  
Db  
2528 ACAAAGTGGCCACATTTTAAAGAAACAGGTGATGCTGGGCCATGAGGATTTTATGAAATGAA 2587  
Qy  
3169 ATGAATCATCTGAAAGACTTCAATATGTAATTTTATCAATTTCTCCAAAACCCAGCATAGCA 3228  
Db  
2588 ATGAATCATCTGAAAGACTTCAATATGTAATTTTATCAATTTTCTAAAACACAGCATAGCA 2647  
Qy  
3229 AGGTGAAATATTCAGAGAAAGGAGAAACCAAGAAATCAAGAAATCAAGAAATCAAGAAATTCGCGC 3288  
Db  
2648 AGGTGAAATATTCAGAGAAAGGAGAAACCAAGAAATCAAGAAATCAAGAAATCAAGAAATTCGCGC 2707  
Qy  
3289 CACCACCACTCTCTGAAATCTGCTGCTGCTTTCAGCAAGTCAATTAACCTTCTCAACTGTG 3348  
Db  
2708 CGCCACCTCTCTCAGAGATTTCTGCTGCTTTCAGCAAGTCAATTAACCTTCTCAACTGTG 2767  
Qy  
3349 ATATCATGATGTAATTTCTCAGGACCGTATTTTGAAGGCGGCAATTAGACAGATTTCTAACT 3408

2768	DB	ATGTTATGATATACATCTCAGACCAATCTTTGTGACGGGCGAGTGGACACCGGAGTCTTAATC	2822
3409	QY	TGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATTTCTGGCAATGGGTTTACTACAAG	3468
2828	DB	TGTGGACAGAAGGGATGCTGCAGATGGGTTCCATATATTTGGCACTGGGCTTGTCTGGAG	2887
3469	QY	AGAAGCAAACAGCTTCAAAAAGCTCCTGAAGAAGTAACAATTTGACTTTTATCATATAAGG	3528
2888	DB	AGAAGCAGCAGCTTCAGAAAAGCTCCTGAAGAGGAAGTGGCTTTTGTGACTTTTACCATTAAG	2947
3529	QY	CTTCAAGATTTGGGAAGTTTCAGCCAT-----GAATATACAATGCTTTTGGAAAAC	3579
2948	DB	CTTCAAGANTTGGGAAGTTTCAGCCATGAATGCTCAGAATATACAATGCTTTTGGAAAGAC	3007
3580	QY	TCAAAGGAATTTCCCGAGTTGAAGGCCCAGAAGGACATGATAACCTGATACTTTCAGATGT	3639
3008	DB	TCAAAGGAAATCCCCCAATTAGAAGGCCAGAAGGACATGATAACATGATACTTCCAGATGT	3067
3640	QY	TTGACACAGTGAAGCGATTAGAAGAAAAATCTTGTTTAAATTTGTAGCAACACATTCAGGAT	3699
3068	DB	TTGACACAGTGAAGCGATTAGAAGAAAAATCTTGTTTAAATTTGTAGTGGCAACACATTCAGGAC	3127
3700	QY	CGGAATCTATTAGAATGATGAGATTACTCATGATAAGAAAAAGCAGACGAAAGAA	3759
3128	DB	TGGAGTGCAATTAAGAGTGAGAGATTACTCATGATAAAGAAAAAGCAGAAACGGAAGAA	3187
3760	QY	AAGCTGAAGCTGTGATGGCTACATCGCCAGAAGATCATGGCTCAGATGCTTGCCTTTACGA	3819
3188	DB	AAGCTGAGGCGCTAGGCTTCATCGCCAGAAGATCATGGCCAGATGCTTGCCTTTACGA	3247
3820	QY	AAAACTCATTTGAAACTCATAACTCATGTATGACAATACATCAGAAATGCTTGGGAAAG	3879
3248	DB	AAAACTTCATTGAAACCCACAACTCATGTATGATAATACTCAGAAAGTAACAGGGAAGG	3307
3880	QY	AAGATTTCCATTATGGAGGAGAGACGCCCCAGCGNGTCAGTGACTACTCTAGAAATGCTT	3939
3308	DB	AAGACTCCATTATGGAGGAGAGACGACCTCAGCAGTCAGTGAGGCGCTCTAGAAATTGCCTC	3367
3940	QY	TGGGTCTTAAACGGGTCCTATCTGTACTGAAAGGAGGTGCTGACGTGCATCCTTTGCC	3999
3368	DB	TGGGCCCTTAAACGGGGCCGCTGTACCGAAAGGAGGTGCTGACGTGCATCCTCTGCC	3427
4000	QY	AAGAAGAACAGGAGGTGAAAATAGAAAATAATGCCATGGTATTAACGGCCCTGTGCCAGA	4059
3428	DB	AAGAAGAACAGAGGTTAAACTAGAAAATAATGCCATGGTATTTGTGACGATGTTGTCAGA	3487
4060	QY	AATCTACTGCTTAAACCAGCACAGGGGAAAAACCATAGAACTCTCAGGAGAGAGCCCTAG	4119
3488	DB	AATCCACGCGCTTAACCCAGCACAGAGGGGAAGCCGTGGGACCACTTAGGGGGAACACTGG	3547
4120	QY	ACCCACTTTTTCATGGATTCGACATTTGGCACTATGGAACCTTATACAGGAAGCTGTGGTCATG	4179
3548	DB	ACCTCTTTTTCATGGATTCGACATTTGGCACTATGGAACCTTATACAGGAAGCTGTGGTCATG	3607
4180	QY	TAATGCACGCACTGTGCTGGCAGAAATATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGC	4239
3608	DB	TAATGCATGCAGTGTGCTGGCAGAAATATTTTGAAGCTGTGACGCTGAGCTCCGACGACG	3667
4240	QY	GCATTCAATGTGACTTTTGTGACTGGAAGTGGAGAAATATCTTTGGCCCTCTTTTGAAGAT	4299
3668	DB	GCATTCACTAGACTCTGTTTGACCTGGAGAGCGCGAGTAGCTATGTCGCCCTCTGCAAGT	3727
4300	QY	CTCTGTGCAATCTGTGATCCCCATATTTCTTTGCAACCTTGAACCTCAAAAGATATAACAGTGAGA	4359
3728	DB	CTCTGTGCAACACTGTCTATCCCCATCATCTTTTGCAGCCGCGAGAAGATCAACAGTGAGA	3787
4360	QY	ATGCAGATGCTCTTGTCTCAACTTTTGAACCTTGCCACCGTGGATACAGACTGTTCTGGCCA	4419
3788	DB	ATGGGAGGCTCTGTCTCACTTTTGACCTTGGCCGCTGGATACAGACTGTCTCTGGCCA	3847
4420	QY	GAATATCAGGTTATAATATAAGACATGCTTAAAGGAGAAAAAC---CAATTCCTATTTTCT	4476
3848	DB	GAATATCGGTTATAATATAAAGCACTCTTAAAGGAGAAAGCCCAAGAGTCTCTGTCTGT	3907

Qy	5545	CAGATGATGAGCGAAAGCATCTGTCTCTCTGCGCTTTTCTGTGGGGCTATATGTTCTC	5604
Db	4988	CAGATGATGAGCGAAAGCATCTGTCTCTCTCTTTTCTGTGGGGCCATCTCTGTGTTCTC	5047
Qy	5605	AGAACATTTTGTCTGCGACGAAATTCGTGAACGGGGAAAGGTTTGGAGCTTGGCATTTTTCACG	5664
Db	5048	AGAACATCTGTGTGCGAAGAAATAGTGAATGGGGAGAGGTTTGGAGCGTGGTTTTTCATG	5107
Qy	5665	CAC TTCACTGTGGAGCGGAGTCTGCATTTTTTCTTAAAAAATCAGAGAAATCCGAGTGGTCC	5724
Db	5108	CGCTTCATTTGTGGTCTGGAGTCTGCATTTTTTCTTAAAAATCGAGAAATCGAGGTGGTCC	5167
Qy	5725	TGGTTGAAGGTAAAGCCAGAGGCTGTGCCTATCCAGCTCTCTTACTTGGATGAATATGGAG	5784
Db	5168	TGGTGGAAAGGAAAGCAGAGGCTGTGCCTACCCAGGCCCTTACTTGGATGAATATGGAG	5227
Qy	5785	AAACAGACCCTTGGGCTCGAAGAGGGGCAACCCCTTTCATTTATCTCGTGGAGCGGTATCGGA	5844
Db	5228	AAACAGACCCAGGGCTAAAGNAGAGAAACCACTTCATTTATCTGGGAGCGGTATCGGA	5287
Qy	5845	AGCTCCATTTGGTCTGGCAACAACA CTGCAATTATAGAAGAGATTTGTAGGAGCCAGAGA	5904
Db	5288	AGCTGCATTTGGTCTGGCAACAGCACTGCATTA TAGAAGAGATTTGTCTGGAGCCAGAGA	5347
Qy	5905	CTAATCAGATCTTAATTTGGATTCAACTGGGAGTTTACTGTGAGCTCCAACTCTGCCTCAAG	5964
Db	5348	CTAATCAGATGCTAATTTTGGATTTTAACCTGGGAGTTTACTCTGAGCTTCTAGTTCTGCCTCAAG	5407
Qy	5965	ACAATCAAAATGACGACAGTAGTAAAGGGCTGATTTCAAAAATATGGHAAACTTTTCTGAGG	6024
Db	5408	ACAATCATGAGTGACATCAATAATAAGACTGATCTAAAATCTTAGAGAACTTTCTGAGG	5467
Qy	6025	GCTGGGAAAGTATTTGGAGGGTCTTTTGGTCCATGTCCAGGTTCATCTTACATCAATAAAT	6084
Db	5468	ACGGGGAAAGTATTTGGAGGGTCTTTTGATCCATGTCCAGATTCAACA CATTAATAAAT	5527
Qy	6085	ATTTC TTAAATGGAGTATTTGCTTTTCAATTTAGCAACATATGCTTCCACAGGAAAAA - AGGAC	6143
Db	5528	ATTCCTTAAATGGAAATATTTGCTTTTCAATATCAAAACATAAGCTTCAAGGHAACAGAC	5587
Qy	6144	ATAGATCAATCTGT-----TTTT	6160
Db	5588	ATAGATTAATGTTTTTATGTTCTAGAACACATAAAGAAATGCTTGTTCATCCAAGTGTCTAT	5647
Qy	6161	ATGTGCTAGTATTTCCAGGAATTTATTTCCCTTCATATAATTGTCTCATTTTCATTTATTTT	6220
Db	5648	TTCTGTCTAATAATTTCCGAAAACTCTTTCCCTTTCATAACTGTCTCTAGTTTCATTTTCATAT	5707
Qy	6221	CATCCACTTGTGTAGATGAAGTCAGTCAAAACAGTTGTAGACATTTTATGTGTGGTTAAAC	6280
Db	5708	CACCACTGTTTAATGAGTCAATTAAGCATTTTGTGGACATTTTCTCCATCTGGCTAAC	5767
Qy	6281	TC TTCTGCAATTTTGTATTTGGGTTTT	6307
Db	5768	ATCTCTGCACCTTTGTATTTGGTGTTT	5794

## RESULT 4

[illegible]









Db 1273 TATTGAGATGCGCTTGGCGGATTTTATGTCAGAGTTGGTTTACAAGAGGCGCCAGATGGT 1332

Qy 1734 GAGAAATCCCTGTCATGAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCC 1793

Db 1333 GAAAACTCTTCTAGTGGGACAGACTGATGCTTAGTGAATCCAAATTTATGGAAAGGTGCT 1392

Qy 1794 CGTAAGATCCCTTCATGAATTCATCTTCAGCAGTTTATTTTATGGAGATGGAATACAAAAA 1853

Db 1393 AGGAGTGATATATCATCAGTGTGTTCAAGGACAGTCTGCTTATGGATTTGAAATACAGAAA 1452

Qy 1854 CTCTTTGCTATGGAATTTGTGAAGTATTTATAACAACTGCGAAGAAATATATCAGTGAT 1913

Db 1453 CTAATTTGCTGTTGCAAAAAATTTACCAGCAGTTGGCAGAGAGATTTTATGGAGGAT 1512

Qy 1914 GATCATGACAGAGTATCTCTATTAACGCACTTTTCAGTTTCAGATGTTTACGTTCTTACT 1973

Db 1513 GATCAGGAGCAGCAGTGTGCGTGACTCTCTATCTGTCCAGTTCTTCACGCGCACCTACT 1572

Qy 1974 CTGGCTCGACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTCATTTACTTGAACCTCTGCTA 2033

Db 1573 CTGGCTCGATGCTCATCAGAGAAACCTTGATGACATATCATTTAAGCTTTTATG 1632

Qy 2034 GAAGTTTT- - - - -ACCTGAGTACTTGGACAGGAACTAATAAATTCACCTTCCAGGGTTAT 2087

Db 1633 GATCAATTTGACATCGAGATGCCAGGGCAGATTTTCAGTTTGAACGATACACTGCTTTA 1692

Qy 2088 AGCCAGGACAAATTTGGGAAGATATATGCAATATATGATGACCTTAAGATATATCTGATC 2147

Db 1693 CAAGCCTTCAAAATTTAGGAGAGTACAGAGCCTTATTTTATGATCTCAAGTATGTTTAAAT 1752

Qy 2148 AGCAAAACCAATATGACAGAAAGATTAAGAAATGCAATGCTTCTTGAAGTTTTCGATCT 2207

Db 1753 AGCAAAACCAATGATGCTGATGAGCTGAGGCGAAGATTTCTTGAAGGGTTTGAATGCC 1812

Qy 2208 TTTTGTGAAGATTTCTTACCTGATGAGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAA 2267

Db 1813 TTTTGTGAATTTACTAAATATGATGACAGGATGGATCCAAATACAGTCAAGTAGGACAA 1872

Qy 2268 CACATTAAGTGGATCTCTGATTTGGAGGCTGCAATGCTATACAGATGCAATTTGAAGAT 2327

Db 1873 CATATGAAATGGAACACAGAGTGGGAGCAGCTTTCACACTACAAATGAAATTAACAT 1932

Qy 2328 ATTTTACTCATGTTCCAGAGTGGTGTCTGATGAAGAACTCTTACTTGTGCTTAT 2387

Db 1933 GTCAATTCATGATGACAGACTGGTGTCTTCCAGATGAAGAAAGTGTAAATCGAGCTTAC 1992

Qy 2388 AAAGAAATGTCACAAAGCTGTGATGAGGTGCAGTACCACTTTTCATATCTAGTAGCAAGACA 2447

Db 1993 AAGAAATGTCGCTGTACTGATGCAAGTGTCAATGCTGTATCTGATGGTGAACAGCCA 2052

Qy 2448 GT- - -AGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCCCTACAGATATCTGAGGAT 2504

Db 2053 ATCACATTAAGCATTTGTGGCAATTCAGTGGAAACATATCAGATATCTGTTTCCCAAGAA 2112

Qy 2505 CTTGTAGCATACATCTGCTCTCTAGGACCTTCTGCTGCTTCTCATGTAGTTTAAAGC 2564

Db 2113 AAAGTTAGCATTCACCTCCAGTTTCTCGCTTACTTCGAGGTTTACATGATATTTAAGC 2172

Qy 2565 AGGCTGGGTGCTGTTTCAAGACTGATGAAATTTGTGCTTTTGTGAGCACTTTCAAGTAGAG 2624

Db 2173 AAAGTGAAGTGGCATATTAATTTCCAGAGCTCCCTACCTTAAGTGAACCTTAGCCACCC 2232

Qy 2625 GTACTAGTGGAAATCTCTTATAGTTGCTGTTGTTGGTGTGCCAGGTTGTTGCTGAGATG 2684

Db 2233 ATGTTGATAGAACACCTCTTAGATATCTTGTGTTCTGCTGTGCGCAAGTACATGCGGGAATG 2292

Qy 2685 TGGCCAGAAATGGAATGCTCTCTTATTAAGCAGGTTTATTAATCAAGATGTTTAAAGTGC 2744

Db 2293 TGGAGAGAAATGGGTCTCTCTAGTAAACAGATTTATTAATCTACATAATGTTGAATGC 2352

Qy 2745 AGAGAAAGAAATGATGATGAAAGATATCATCTGCTTTCAGATTTGGTGCATCTTTTATGGAT 2804

Db 2353 AGACGTGAGATGTTTGAACAAGATGATGTAATGCTTTCAGACAGGTTGCTCCATGATGAT 2412

Qy 2805 CCCAATAGTTTCTTGTACTGTTACTTTCAGAGGTATGAACCTTTCGCGAGGCTTTTAAACAAG 2864

Db 2413 CCAAAATCATTTCTTGATGATCATGCTCAGCGCTTTTGAACCTTTATCAGATTTTCAGTACT 2472

Qy 2865 ACCATATCTACAAA- - - - -GACCAGGATTTTGAATTAACAATAT 2903

Db 2473 CCAGACTATGGAAAAAGATTTAGTTCTGAGATTACCCATAAGGATGTTGTTTCAGACAAG 2532

Qy 2904 AATACATTAATAGAAAGAAATGCTTCAGGTCTCTCATCTATATTTGTTGGGTGAGCGTTATGTA 2963

Db 2533 AATACTCTTAATAGAAAGAAATGCTATACCTCATTAATAATGCTTGTGAGAGAGATTTAGT 2592

Qy 2964 CTTGGAGTGGGAATGAGCCAAAGAGGTCAATAGAGAGAAATCATTCACCTTGCTT 3023

Db 2593 CTTGGAGTGGAGCAGGTAATGCTACAGATGAATCAAGCGAGAGATTTATCCATCAGTTG 2652

Qy 3024 TGCAATGAACCCATGSCACACAGTCACATGCCCCAAAATTTTACCTGAGAAATGAATAAT 3083

Db 2653 AGTATCAAGCCTATGGCTCATAGTGAATTTGTAAGTCTTTTACCTGAAGATGAGAACAA 2712

Qy 3084 GAAATCGCTTAGAGATGCTATPAACAAGATGGCCACATTTTAAGAAACACAGGTGTATCA 3143

Db 2713 GAGACTGGCATGGAGAGTGAATCGAAGCAGTTGCCCATTTTCAAGAAACCTTGGATTAACA 2772

Qy 3144 GGCCATGAGTTTATGAACCTAAAGATGAATCACTGAAAGACTTCAATATGTAATTTTAT 3203

Db 2773 GGACGAGCATGTTATGAACCTGAACCAAGATGTCGCAAGAGATTTCACTTGTATTTCTAT 2832

Qy 3204 CATTACTCCAAAAACCCAGCATAGCAAGCTGAAATATATGCAAGAAAGAGAGAAAAACA 3263

Db 2833 CACTTTTCAAGGGCAGAACAGTCCCAAGCAGAGAGAGCGCAACCGAAAATTTGAAAGACAA 2892

Qy 3264 GAAAAAAGATGAAGCAATTTGCCGCCACCACTCCTGTAATTTCTGCTGCTTTTCAAGC 3323

Db 2893 AATAGAGAAGATACAGCACTCCCACTCCGCTGTTGCTCCATTTTGCCTCTCTGTTGCA 2952

Qy 3324 AAAGTGATTAACCTTCTCACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAG 3383

Db 2953 AGCTGGTTAACTATTTTGGAGTCAAGTGTCAATGTTGTCATCATGGGAACTTTCTGCA 3012

Qy 3384 CGGCAATATAGACACAGATTTCTAACTTGTGGACCGAAGGATGCTTCCAAATGGCTTTTCA 3443

Db 3013 TGGGCTGTGGAACATAATGGATATGCTGCTGCTCAGAGTCCATGCTGCAAGAGGTTTACAT 3072

Qy 3444 ATTCTGGCATTTGGGTTTACTAGAAAGAGACCAAGCTTCAAAAAGCTTCTGAG- - -AA 3500

Db 3073 TTAATTTGGCATGGCACTACAAAGAAAGAAAAACAATTTAGAGAAATGTCAAGAAAGAGCAT 3132

Qy 3501 GAAGTAACTTTGACTTTTATCATAGGCTTCAAGATTTGGGAAGTTTCAAGCCATGAAT- - - 3557

Db 3133 GTAGTAACTTTACCTTCACTCAGAGATATCAAACTGTGTGAGCGCCCAAAAATTTCT 3192

Qy 3558 - - - - -ATACAAATGCTTTTGGAAAAAATCTCAAGGAAATTTCCAGTTTAGAAGGCGCAAG 3611

Db 3193 CCTAGCATACTAGCTATGCTGGAACACTACAAAATGCTCCCTACCTAGAAAGTCCACAAA 3252

Qy 3612 GACATGATTAAGTGGATCTTCAGATGTTTGAACAGTGAAGCGGATTAAGAGAAAAATCT 3671

Db 3253 GACATGATTCGTTGGATATTTGAAGACTTTTAATGCTGTTTAAAGATGAGGAGAGTTTCA 3312

Qy 3672 TGTTTAAATTTAGAACCACTCAGGATCGGAATCTTATTAAGATGATGAGATTTACTCAT 3731

Db 3313 CCTACCAGTC- - - - -CCGTGGCAGAGACAGAGGAACCAATTAAGAGAGAGTTCAAGG 3366

Qy 3732 GATAAGAAAAAGCAGAAACGAAAAAGAAAGCTTGAAGCTGTAGGCTACATCGCCAGAG 3791

Db 3367 GACAAAGCAAGCTGAGAGGAGAGAAAAACAGAGATTTGCCAGACTGCGCAGAGAGAAAG 3426

Qy 3792 ATCATGGCTCAGATGTTGCTTACAGAAAACTTTCATTGAACTCATATAAATCATGTAT 3851

Db 3427 ATCATGGCTCAGATGTTCTGAAATGCGCGCATTTTATTTGTAAGAAAAACAAAGAACTCTTT 3486

3852 GACATATACATGAAATGCTGGGAAAGAGATTCATATATGGAGGAGAGGACCCCA 3911  
3877 CAGCAGACATAGAACTGGATGCTCAACCTCTGCTGTTCTTGATCA-----TAGCCCT 3540  
3912 GCAGTCAGTACTACTAGAAATGCTTGGGTCTTAAACGGGTCCATCTGTTACTGAA 3971  
3541 GTGGCTTCAGATATGACATTAACAGACTGGGTCCACACAACTCAGGTTCTGACAA 3600  
3972 AAGGAGGTGCTGACGTGCACTCTTTGCCAAGAAAGACAGGAGTGAATAAGAAATAAT 4031  
3601 AGACAATTCGTTACATGATATGTTGTCAGAGGAGCAAGAGTTAAAGTGGAAAGCAGG 3660  
4032 GCCATGTTATATCGGCTGCTGTCAGAAATCTACTGCTTAAACCCAGCAGAGGGA 4091  
3661 GCAATGGTCTTGGCAGCATTTGTTGAGAGATCACTGTTATTAACAAAGACAGAGTAA 3720  
4092 CCCATAGAACTCTCAGGAGAGCCCTAGACCCACTTTTCATGATCTCAGACTTGGCATAT 4151  
3721 TTTATTCAGATCCAGAAA-----TATGATCAATTAATCATGACCCCTGATCTGCTGT 3777  
4152 GGAATCTTATACAGGAGTGTGTCATGTATGACGAGTGTGCTGCGAGAGTATTTT 4211  
3778 GGAACACACTAGTAGTGTGGGCACATATGCAATGCCATGTTGTCGCAAGGTATTTT 3837  
4212 GAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTTT-----T 4259  
3838 GATTCGCTCAAGCTTAAAGAACAGCGAAGGCAACAGAGATTACGCTTACATACGAGCTAT 3897  
4260 GACTTGGAAAGTGGAGATATCTTTGCCCTCTTTGCAATCTCTGCAATCTGATC 4319  
3898 GATGTAGAAACGGAGAAATCTTTGCCCTTTTGTGAATGCTTGAGTAATCTGTTAT 3957  
4320 CCCATTAATTCCTTTCACCTCAAGAGATAAACAGTGAAGATGACAGATGCTCTTCTCAA 4379  
3958 CC-----TCTGCTCTCTCCAGAAATATTTTAAACAGGTAAATTTTTCAGAC 4011  
4380 CTTTTCAGCTGCGGACGCTGATACAGACTGTTCTGGCCAGAAATACAGGTTATATATA 4439  
4012 CAACAAATCTGACTCAGTGGATTAAGAACAAATATCTCAGCAAAATAAAGCATTA 4071  
4440 AGACATGCTAAGGAGAAACCCATTCCTATTTCTTTTAAATCAAGGATGGAGATTCT 4499  
4072 CTTAGGAAGAGAAAGTACTCTTAATATGCTCTACAAGAAATTCAGAAATGTGAT 4131  
4500 ACTTTGAGTTCCATTCATCTGAGTTTGGCGTTGAGTCTTCGATTAATAATTAATCAAT 4559  
4132 GAAATACAGCTCCCTGAAGGTTGAGGCTGATTTCTGCTCTAGATCCCTTATCTGAG 4191  
4560 AGCATCAAGGAAATGTTATCTCTTTGCCACAAATTTATAGAAATGGAATTTGAAAGTG 4619  
4192 AGCATAAAGAAATGCTTAAACGACATTTGGAACTGCTACTACAGGTGGGACTAAAGGT 4251  
4620 CCACTGATGAAGGATCTCGAGTCCCATGCTGACCTGGAGCACCTGGCTTCACT 4679  
4252 CATCCAAATGAAGAGATCTCGTGTCCCAATATGTTGGGTGAGCTGGCGTACACC 4311  
4680 ATCCAGGCAATGAAATCTATTTGGAGATGAAGGAAACCTCTGTTTGGAGCACTTCAA 4739  
4312 ATCCAAAGCATAGAAAGAAATTTGATGATGAGATTAACCAATGTTTGGTCTTACT 4371  
4740 AATAGGACGCAATATGTTGAAAGCATTAATGACAGTTTGGAGTTGCAAGAGGATACC 4799  
4372 TGCACTGAGTACTGCTTAGGTCAATTGACGAGATTTGGCGCAGCACCTGGACAGTG 4431  
4800 TGTCTCAGTCTGATACAGAAACATCTGTTCTGTTCTTCTATCAGTTGTTCTTCTAAC 4859  
4432 GCATCAGTTTCAAGTGTGCAAGGACATTTTGTAAACTTTTGTGATCACTGGTGCCTAAT 4491  
4860 ATAAATCAGAAATACACCAATCTCTGCTATATAGATCTGTTTCAATGTTTGGTGGT 4919  
4492 GACAGCCATGAGAACTTCCATGCAATATAGATATTTGACATGTTTCAATTTATGTTGGC 4551  
4920 GCTGTGTTAGCATTCCTTCTGTTATTTGGGATGACCCCTGTTGATCTGACGCTTCTTCA 4979

4552 TTGGTCTTCATTTCTGCGTTCAGTG-----TCAGATTTTTCAGGG 4596  
4980 GTTAGTTCTTCTTATAACCACTTTTATCTTTCATTTGATCACCATGGCACATGCTT 5039  
4597 ATCAGCTTGGCACTGGAGACTTTCACATTTTCATCTGGTTACTATGGCACATCATA 4656  
5040 CAGATATCTATTAACAGTATAGACACAGGCC-----TACCCCTTGTCTCAGTTTCAAGAAGACAGT 5096  
4657 CAGATCTTACTTACCTCATGTACAGAGAGAAATGGCATGGATCAAGAAAAATCCCCCTTGT 4716  
5097 GAAGAGGCTCATTCGGCATCTTCTTCTTTCGAGAAATTTCTCAATATACAAAGTGGCTCC 5156  
4717 GAAGAAGAAATCAGCAGTTCTTGTGTTATATAAACACTTTCACCATATACGGGAAGTGC 4776  
5157 ATTGGGTGTGATATTCCTGCTGCTGTTTGGGTCTCACTGAAGAAATGGCATCACCCCT 5216  
4777 TTGAAGAAATACCATCCGCTGCACTCTGGAGGAGTGTGAGAGCTGGATCATGCT 4836  
5217 TATCTGCTGCTGCTGCAATGTTTTTCACTATTTTCTGGGTTAACTCCGCTGAGAA 5276  
4837 TTCTGAGTGTCTGCTTATTTTCTTAAATGAGTTCTTCCACCCGAC 4896  
5277 CTGCATACCAATCTGCAAGAGGAGATACAGTGCATCTGCTGATATCTATCTTACT 5336  
4897 ATTCA-----AGTTCTTGGAAACAGCCATTTTGAACTATTTATGTAGCTATCTTCCCTACCA 4953  
5337 ACAATTTGTTCTGCTCTTCCAGGAATATTTGGGATCTGTAAGGCCCTTGTCTCAGAGG 5396  
4954 AACAACTCTATTTGCCCTTTTCAAGAAATATGATGATTAATGATTTCTGATGAAAT 5013  
5397 TGGTGTGAGATCTGCTCTTAACTTAACTGTTTGAAGCAAAAAACACCGTGGTCAAGTAC 5456  
5014 TGGTGCCGTAACAGTGAAGTTAAAGATATCTAGAGGTGAAAGAGATGCTATAAGATAT 5073  
5457 CTTAGAAAGAAATATGTTGATAGAGCTTCTGATGATCTGCTGCTCTGATCA 5516  
5074 CCNAGAGAACTTAACTTAACTTCCAGAGGATTAACAGAGCTCATTAATCAA 5133  
5517 GCTTCTCATTTTCAAGTCCCACTGCTGATGATGATGAGCAAGATCTGCTCTGCTG 5576  
5134 GCATCCAAATTTCTGCTGCTGCGGAAATCAGTGGTGTATAGAGCAGAGCCCACTCTGTC 5193  
5577 TTTTCTGCTGCTGCTATCTATGTTCTCAGAACTTTGCTGCCAGAAATTTGTGAACGG 5636  
5194 CTTGTGCGGATCTCTGCTGCTGCTCCAGAGTTTACTGCTGCCAGACTGAACCTGGAAGG 5253  
5637 GAAGAGTGGAGCTTGCATTTTTCACGCACTTCTGCTGAGCCGAGTCTGCAATTTTC 5696  
5254 GAGGATGTAGAGCTTGCACAGCTCACACCTTCTGCTGCTGCTGAGTGGGCTCTTTC 5313  
5697 CTAAAAATCAGAAATGCGAGTCTGCTGTTGAAGTAAAGCCAGAGGCTGTGCTCTAT 5756  
5314 CTGAGATACGGAAATGTCAGTGTCTATTTTGTGCAAAACCAAGGCTGTTTTTAT 5373  
5757 CCAGCTCTTACTTGGATGATATGAGAAACAGACCTGCTGCTGAGAGGGGCAACCCC 5816  
5374 TCTCTCTTACTTGTGATGATATGAGGAGACCGGAGCTCAGACGGGAAATCCT 5433  
5817 CTTCAATTTCTGAGCGGTATCGGAGCTTCAATTTGCTGCTGCAACCACTGCAAT 5876  
5434 TTCAATTTATGCAAGAGCGATTCAGAAAGATTCAGAAAGCTCTGGCACCAACAGATGTC 5493  
5877 ATAGAAGAGATTTGCTAGGAGCCAAAGACTAATCAGATGTTTATTTTGGATTCACCTGGCA 5935  
5494 ACAGAGAAATTTGACATGACAGAAACCAATCAGACACTGTTGGATTTGACTGGCA 5552

RESULT 6

US-09-724-126A-5

; Sequence 5, Application US/09724126A

; Patent No. 6706505

; GENERAL INFORMATION:

; APPLICANT: Han, Hui-Quan  
; APPLICANT: Kwak, Keith  
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family  
; FILE REFERENCE: 01017735966A  
; CURRENT APPLICATION NUMBER: US/09/724.126A  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,211  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 6089  
; TYPE: DNA  
; ORGANISM: Moube  
; US-09-724-126A-5

Query Match 19.6%; Score 1233.6; DB 4; Length 6089;  
Best Local Similarity 54.7%; Pred. No. 0;  
Matches 2870; Conservative 0; Mismatches 2269; Indels 105; Gaps 17;  
  
QY 769 TGGGATCTTGGTGGGATCAGCAAGTTGATTTTATACCTGCTTCTTGATCATTTGGCAC 828  
DB 836 TCGCAGGAGATGGTGCAGCAACCGACCTCAACAGAGAAAGTGATACAGCAATTTAGCCC 895  
QY 829 AATTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAA 888  
DB 896 ACTGTGTGCCMAATCTACTTGCCTGGGGCCCTAACCCCTTCCCTCAGAAAGAGACAGC 955  
QY 889 GTGTACAAATGTCAATATTTCACTCACTGGAAATGGTACTTATTTTGGAGAGATCCAGATA 948  
DB 956 TGGCACACACATCTGCTGGGACCGATGGAGTGTGTACATCTGCGCTGAAGACCTGGCG 1015  
QY 949 TTTGCTTAGAGAAATTTGAAGCAC---AGTGAGCAATTTACAGTTTGTGGAGGGTTTCA 1005  
DB 1016 TGGGATTTCCAAAGCTCGAGAGGCAACAAAGCCTTCTCACCTCTGTGGCGAGTGTTA 1075  
QY 1006 AAAGTGGAGAGACACCTATTTCTGACGGGATTTGCAATTTGATCCACATGTGTACTCT 1065  
DB 1076 AAGTGGGGGAACCTTACATCTCTGACAGACTGTGCGAGTTGACCCCACTGTGTAT 1135  
QY 1066 GTATGGAATCTTCCAGGACAGTGTTCATAAAATCATCGTTTACAGATGCATCTCTA 1125  
DB 1136 GCATGGAGTGTCTCTGGGAGATTCATAGACCATCGATATAGGATGACCATCGG 1195  
QY 1126 CTGGAGGAGGGTCTGTGATCTGTGGAGACACAGAGGCAATGAAACTGGCCCTTTTGTG 1185  
DB 1196 GAGGAGGGGGTCTGTGACTGTGTGACACTGAGGCGTGGAAAGAGGGACCTTACTGCC 1255  
QY 1186 TAAATCATGAACCTCGGAAG---AGCAGGTACTATAAAGAGAAATTCACGCTGTCCGTTGA 1242  
DB 1256 AGAAGCAACAGCTCAGCAGCTCTGAAGTTTGGAGGAGAGAGATCCTCTTGTGCATCTAT 1315  
QY 1243 ATGAAGAGTAATTTGCAAGCAGGAAATAATTTCTTCAAGTGTATAAAATATGTCTGATG 1302  
DB 1316 CAGAAGATGTATCGCCAGACTTACACATTTTGTCTATTATTTTTCGATATGCGATG 1375  
QY 1303 AAATGACTATATGGAGAGGAGAAAGAACTGCTCTGAACTCCAGATCAAGGAGAGAAA 1362  
DB 1376 ATATACCTCTGGGAAAGAAAGTGAATTTGCTTGAAGACTTAGAAGTGGCAGAGAGA 1435  
QY 1363 ATGAAGATATCTATTGTGCTCTTTTCAATGATGAACACCATTTCAATGACCAAGCTCATAT 1422  
DB 1436 GTGACACCTTACTGCTATGCTGTGTTTATGATGAGGTTTCAACCTTATGAGCAAGTCA 1495  
QY 1423 ACAGGCTTCAAGAGCTCTTGAATGTGAGCTGCGAGAGGCCAGTTTGCATACCACTGCCA 1482  
DB 1496 ATACCTTTCAGAAAGCTGTGAATCTGTACAGAGAGAGGACCATTTGGCTTTTGCATCAG 1555  
QY 1483 TTGACAAAGAGGGTCTGTGGGCTGTGTTAAAGCGGAGCTTATGCTGCTTCCAGGAGCAA 1542  
DB 1556 TTGATCGAGATGSCGCTGAGGCTGTCCGATATGGAGATTTCCAGTACTGTGTATCAAGCA 1615  
QY 1543 AGGAAGATATAAGAGTCAATTCAGAAATGTCTCTCAACATCCACTTCAATGTAGAAGTAT 1602

DB 1616 AGACAGTCAATTTGTGAGGAACACACAGCAGACAGACCAAG---CCGCTCAAAGTTCAAAGTTA 1672  
QY 1603 TACACTCAGAGATATATGGCTCATCAGAAATTTGCTTGGCTTGTGGTCTCTGGATGAACA 1662  
DB 1673 TGCATCTCCGTGGCTGCTCATCAGAAATTTGGTTTGAAGCTCTCTCGTGGCTGGAA 1732  
QY 1663 AAATTATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCAGAGCATGCTCTAGAGAG 1722  
DB 1733 GTGTTATTTGATCTCAGATGGCCCTTCGAGGATTTTGTCAAGTTGGATTACAGAGAG 1792  
QY 1723 AACCTGACTCGGAGATCCCTGTCTCATTAAGCAGGTTTAATGCTTTTGGGATGCAAGCTTT 1782  
DB 1793 GTCCAGATGGCAAACTCTTCTCTGCTCGACAGACTGATGCTTAATGATTTCCAAATTA 1852  
QY 1783 ATAAAGTGGCCGTAAGATCTTTCATGATTTGATCTTCAGCAGTTTATTTTATGAGATGG 1842  
DB 1853 GGAAGGGGCTAGGAGTGTGTATCACAGTTGTTTCATGAGCAGCTGCTCATGACCTCA 1912  
QY 1843 AATCAAAAAAATCTTTTGTGTAAGTATTAATAAACAATGCAAGCACTGCGAGAAAGAT 1902  
DB 1913 AGTATAAGAGCTGTTTCGCTTCAATTTGCTAATAAATACCGCAGTTTGCAGAGGAT 1972  
QY 1903 ATATCAGTGTATGATCATGACAGAAAGTATCTCTATACTGCACTTTTCAAGTTTCAAGTTTA 1962  
DB 1973 TTATGGAGGATGATCAAGCGGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2032  
QY 1963 CTGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2022  
DB 2033 CCGCACCGCCTGCGCGAATGCTCTCTCAAGAGAGAACCTGATGACCGCTTATCAATTA 2092  
QY 2023 AAATCTGCTAGAGTTT-----ACCTGAGTACTTTGGACAGGAAACAATAAATTTCAACT 2076  
DB 2093 AGGCTTTTCAAGCAATTTGAAACAAGAGATGCCAGGCGAGATTCAGTTTGAACGCT 2152  
QY 2077 TCCAGGGTTATPAGCCAGGACAAATTTGGGAAAGATATATGCAAGTAAATATGACCTTAAAGT 2136  
DB 2153 ACATGCCCCCAAGCCTTCAAGTTTCAGGAGTCCAGAGCCTCATCTTAGATCTCAAGT 2212  
QY 2137 ATATCTGATCAGCAAAACCAATATGACAGAAAGATTAAGAAATGAGTTTCTTGTGAG 2196  
DB 2213 ATGTATTTGATAGCAAAACCAACGAGTGGTGCAGATGAGCTGAGGCGAGAGTTTCTTACAG 2272  
QY 2197 GTTTTCGATCTTTTTCGAAGTTCTTACCTGTAAGAGGAAATGCAAGGAAATCCGAAGAC 2256  
DB 2273 GGTTCGATGCTTTTGTGAATTTCTGAAGTCAAGAGGAAATGCAAGGAAATCCGCTC 2332  
QY 2257 AGTTTGGGCAACACATTTGAAGTGGATCTGATTTGGGAGGCTGCCATTTGCTATACAGATGC 2316  
DB 2333 AGGTGGGACAGCAATTTGAGATGAGCCAGAGTGGGAAAGCAGCTTCACTGCGAGATGA 2392  
QY 2317 AATTGAAGATATTTTACTCATGTTCCAGAGTGGTGTGCTTGTGATGAGAACTCTTAC 2376  
DB 2393 AGCTGACACACGTCATCTCAATGGTGCAGGACTGGTGTGCTCTGAGCAAAAGGTTTAA 2452  
QY 2377 TTGTGGCTTATAAGAAATGTCACAAAGCTGTGATGAGTGCAGTACCAAGTTTCTATCTA 2436  
DB 2453 TTGAAGCTTACAGAAATGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2512  
QY 2437 GTAGCAAGAGACT---AGTACAATCGTGTGGACATAGTTTGGAAACAAAGTCTTACAGAG 2493  
DB 2513 GTGAACAGCAATCACACTCAGTATTTGTGACACTCGGTGGAAACCACTCAGATCTGTG 2572  
QY 2494 TATCTGAGGATCTTGTGAAGCATATCTGCGCACTCTCTAGAGACCTTGTGCTGCTGCTGCTGCTGCT 2553  
DB 2573 TTTCCCAAGAAAAAGTTAGCAATTCACCTCCCAATTTCTCGCTTGTGCTGCTGCTGCTGCTGCTGCT 2632  
QY 2554 TACGTTTAAAGCAGGCTGGTGTCTGTTTCAAGACTGCAATTTGTGCTCTTTTTCAGGACT 2613  
DB 2633 TATTGTTAAGCAAAAGTGAAGTGCATATAAATTTCCAGAGCTCTTACCTCTAAGTGAAC 2692  
QY 2614 TTCAAGTAGAGGTACTAGTGGAAATCTCTTTTACGTTGTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2673

Db 2693 TGAGCCACCATGTTGTAGAACATCTCTTAGATGCTTCTTATGTGCTCAAGTGC 2752  
Qy 2674 TTGCTGAGATGGCGAAGAAATGACATGCTCTTATTTAGCCAGGTGTTTATACCAAG 2733  
Db 2753 ATGCTGGATGTGGAGAAATATGCTCTCTAGTAAATCAGATCTATTAACCATTA 2812  
Qy 2734 ATGTTAAGTCAGAGAGAAATGATGATAAGATATCATCATCTTCAGATTTGTGCAT 2793  
Db 2813 ATGTGAATTCAGGCGAGAGATGTTCTCAAGGACATAGTATGCTTCAGACAGGTGCT 2872  
Qy 2794 CTTTAAATGATCCCAATAGTTCTTGTACTGTGTAATCTCAGAGGTATGAATCTGCGAGG 2853  
Db 2873 CCATGATGGACCCAAACCACTCTCTGATGATCATGCTCAGCGCTTTGAACTCTATCAGC 2932  
Qy 2854 CTTTAAACAGAACCAATATCTCAAAAG-----ACCAAGATTTGA 2892  
Db 2933 TCTTCAGCACGCTGACTATGGAAGAGATTCAGTTCTGAGGTATCCCATAGGACGTCG 2992  
Qy 2893 TTAACAATATAATACACTAATAGAGAAATGCTTCAGGTCTCATCTATTTATGTTGGTG 2952  
Db 2993 TTCAGCAGAAACAACACTCTGATCGAAGAGATGCTCTACCTCATCATCATGCTTTGTGGAG 3052  
Qy 2953 AGCGTTATGTACTCGAGTGGGAAATGTGACCAAAAGAGAGTCAAAATGAGAGAAATCA 3012  
Db 3053 AAGAATTCACCTGGGTTGGACAGTGGCTGCCACAGATGAATCAAGAGGGAGATTA 3112  
Qy 3013 TTCACTGCTTTGCAATGCAACCCATGCCCCACAGTGCCATTCGCAAAATTTACCTGAGA 3072  
Db 3113 TCATCAGTTGAGCATCAAGCTATGGCTCAAGTGAGTGGTGAAGTCTCTGCTGAAG 3172  
Qy 3073 ATGAAATTAATGAATCTGCTTAGAGATGTGATTAACAAGAGTGGCCACATTTAAGAAAC 3132  
Db 3173 ATGAAACAAGAGAGACCGCATGGAGAGCGTCAATCGAGTCCGTTGCAATTTCAAGAAAC 3232  
Qy 3133 CAGGTGATCAGGCCATGAGTTTATGAATGAATGAATGAATGAATGAATGAATGAATGA 3192  
Db 3233 CTGGCTCAGCGGCGAGCATGTATGAGCTGAGCGAGTGGCCAAAGTTCAACC 3292  
Qy 3193 TGTACTTTTATCATTTACTCCAAACCCAGCATAGCAAGGCTGAAATATGCAAGAGAA 3252  
Db 3293 TGTATTTTATCATTTCTCCAGGCGAGCAGTCCAAAGCAGAGAGTCAAGCGAAAT 3352  
Qy 3253 GGAGAAACAGAGAAACAAAGATGAAGCATTTGCCGCCACCACTCTGTAATTTCTGCC 3312  
Db 3353 TGAAGAAGAGAAATTAAGAGATACAGCATCTCCCTCTCCGCTTTTGCCACCGTTCTGCC 3412  
Qy 3313 CTGCTTTCAGCAAGATGATTAACCTTCTCACTGATATCATGATGATGATGATGATGATG 3372  
Db 3413 CTTTGTTCGCGAGTCTGGTTAAACATCTTGCAGTGTGACGTATGCTGTACATCATGGGAA 3472  
Qy 3373 CCGTATTTGAGCGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGCTCCAAA 3432  
Db 3473 CGATCTCGAGTGGCTGTAGACATCAGGGTCTGCTGTGTCAGAGTCCATGCTACAGA 3532  
Qy 3433 TGGCTTTTATATCTCGCATTTGGTTTATAGAAAGAGAGCAACAGCTTCAAAAGCTC 3492  
Db 3533 GGGTCTGCTATTTGATCGGGATGGCTCTCCAGGAAGAGAGCAACCTTGGAGAACGCCG 3592  
Qy 3493 CTGAGAGAGAGT-----AACATTTGATTTTATCATTAAGCTTCAAGATTGGGAAGTTGAG 3549  
Db 3593 TGGAAAGGCGACGTGACAGACCTTCACCTTCACACAGAGATTTTCAAGCCTGTGTATGCAC 3652  
Qy 3550 -----CATGAATATACAAATGCTTTTGGAAATACTCAAGGAATTCGCCAGTTAG 3600  
Db 3653 CACATACTCCCGAGCATCTAGCTATGCTGGAGACCTTGCAGAACGCCCTCTCCCTGG 3712  
Qy 3601 AAGCCAGAGAGACATGATTAAGTGGATATCTTCAGATGTTTGAACAGAGTGAAGCGATTA 3660  
Db 3713 AAGCCACAGAGACATGATGATGCTGCTTAAAGATGTTTAAATGAAGATGA 3772  
Qy 3661 GAGAAATCTTGTTTAAATTTAGCAACCAATCAGGATCGGAATCTATTAAGATGATG 3720  
Db 3773 GAGAGTGTTCACAGACCGCTGTGGCC-----GAGCGGAGGGAACCATTAATGGAGG 3826

Qy 3721 AGATTACTCATGATTAAGAAAGACAGAAACGAAAGAAAGAAAGCTGAAAGCTGTAGGCTAC 3780  
Db 3827 AGAGCTCAAGAGACAAGGACAAAGACAGAGAGGAAAGAAAGCCAGATCGCCAGACTGC 3886  
Qy 3781 ATGCCCAAGATCATATGCTCAGATGCTGCTCTACAGAAACATTTCAATGAAATCTATA 3840  
Db 3887 GCCGGGAAGATCATGCCCCAGATGCTGAGATGACGGCAGCTTCATTGACGAAACA 3946  
Qy 3841 AACTCATGTATGACAAATACATCAGAAATGCTCGGAAAGAAAGATTCCTATTATGAGGAAG 3900  
Db 3947 AAGAGCTCTCCAGCAGACCTTAGAGCTGGACACCTCTGCTCTGCCACTCTTGAC----- 4002  
Qy 3901 AGAGACCCCAAGCAGTCACTACTCTAGAAATGCTTTGGGTCTTAAACGGGGTCCAT 3960  
Db 4003 --AGCAGCCCTCCGCTTTTCAGACGAGCTCTTACAGCACTGGGCCAGCAGACACAGG 4060  
Qy 3961 CTGTACTGAAAGAGGAGTGTGACGTCATCTCTTTCGCAAGAGAAACAGGAGGTGAAA 4020  
Db 4061 TCCCTGAACCCGAGACAGTTTGTCACTGTATATATGTCAAGAGAGCAAGAGGTGACTG 4120  
Qy 4021 TAGAAAATAATGCTCATGTTATTCGGCTGTGTCAGAAATCTACTGCTTTAACCCAGC 4080  
Db 4121 TGGGAGCAGGCGAGTGTCTTGGCAGGTTTGTTCAGAGGTCAACGGTCTGTCAAAG 4180  
Qy 4081 ACAGGGGAAACCCATAGAACTCTCAGAGAGGCCCTAGACCCTTTTCATGAGATCCAG 4140  
Db 4181 ACAGGACGAAACCATCGCGACCCAGAAATAAT---GATCCATTAATTCATGACCCCG 4237  
Qy 4141 ACTTGGCATATGGAATCTTATACAGGAGCTGTGTCATGTAATGCAAGCAGTGTCTGGC 4200  
Db 4238 ATCTGTCTGTGGGACACACACTGGCAGCTGTGGCAGCTTATGCAATGCCATTTGTTGGC 4297  
Qy 4201 AGAAGTATTTGAAGCTGTAC-----AGCTGAGCTCTCAGCAGCGCATTCATGTTG 4251  
Db 4298 AAAGTATTTGATTCGTTCAAGCAAGGAGCAGGAGGAGCGGCTGCGCTGCG 4357  
Qy 4252 ACCTT---TTTGACTTGGAAAGTGGAGAAATATCTTTGCCCCTCTTTGCAAAATCTCTGTGCA 4308  
Db 4358 ACATAGCTAGATGATAGAAATGCGAGTTCCTCTGCCCCGCTCTGTGAGTGCCTGAGCA 4417  
Qy 4309 ATACTGTGATCCCATTTATCTTTTGCNAACCTCAAGATATAACAGTGAATGACAGATG 4368  
Db 4418 ACACGGTGAATCC---CTGCTGCTTCTCCAGGAGCATCTCAGCAGGAGGTTAA 4471  
Qy 4369 CTCTGTCTCAACTTTTGACCTGCGACGCTGATACAGATGTTCTGCGCAGAAATATCAG 4428  
Db 4472 ATTTTCAGACCAACAGATCTGGCAAGTGGACGAGGAGTAAACAGCAGATTAAGG 4531  
Qy 4429 GTTATAATATAGACATGCTTAAAGGAGAAACCCCAATTTCTTATTTTAAATCAAGGAA 4488  
Db 4532 TGGTCCAGATGCTGAGGAGAAAGCAATGCTGTGACACGCTCTCTTCAGAGGACACAG 4591  
Qy 4489 TGGGAGATTTCTATCTTTGGAGTTCCATTCCTGAGTTTGGCGTTGAGTCTTCGATTA 4548  
Db 4592 AAGCCATGAATATATACCTATCCCGAAGGCTTCAGGCTGATTTTATCTTAGGAAACC 4651  
Qy 4549 AATATTCAAATAGCATCAAGGAATGTTATCTCTTTGGCCACAACAAATTTATAGAAATG 4608  
Db 4652 CATATCTGATAGATTAAGAAATGTTAAACAAATTTGGAAACGCTCTTACAGGTGG 4711  
Qy 4609 GATTGAAAGTCCCACTGATGAAAGGATCTCCGAGTCCCATGCTGACCTGGAGCACT 4668  
Db 4712 GACTGAAGTTTCACTCTAAATGAAGTGAACCCCGTGTGCCCATCTCTGTGCTGGGGACCT 4771  
Qy 4669 GCGCTTTCTATCTCAGGCAATGGAATCTATTTGGGAGATGAAGGAAACCTCTGTTTGG 4728  
Db 4772 GTGCATACCATCCAGAGCATAGAAAGAAATTTTGAAGTGAAGAGAGAGGCTGTTTTG 4831  
Qy 4729 GAGCATTCAAAATAGGAGCATATGCTGTGAAGCATTAATGACGTTTCAGCTTGCAC 4788  
Db 4832 GACCTCTGCTTGTAGACTGAGACGCTGTCTCAGGTGCTTAAACAGCGTTTGCAGGACAC 4891



4789 AGAGGATTACCTGCTCCTCAGGTCCCTGATACAGAAACATCTGGTCTGCTCTATCAGTTG 4848  
|||||  
4892 ATTGACAGTGGCGTTACTTCTGTGGTACAAGGACACTTCTGTAAACTCTTTGCATCCT 4951  
|||||  
4849 TTCTTCCCTAACATAAAATCAGAAGATACACCATGCTTCTGTCTATAGATCTGTTTTCATG 4908  
|||||  
4952 TGGTGCCTAGTGACAGACTATGAAGACCTCCCGTGCACTAGACATCGACATGTTTCACT 5011  
|||||  
4909 TTTTGGTGGGTGCTGTGTAGCAATCCCATCTTGTATATGGGATGACCTGTTGATCTGC 4968  
|||||  
5012 TGCTGGTGGGCGTGTGTGCTGCTTCTTCCAGCTCTGCACTG-----TCAGG 5056  
|||||  
4969 AGCCTTCTTCAGTGTAGTTCTTCTCTATAACCACTTTATCTCTTCCATTTGATCACCATGG 5028  
|||||  
5057 ATTTTTCAGGAAGAGCTCTGGGCACTGGGACCTGCACATCTTCACTTGGTTACCATGG 5116  
|||||  
5029 CACACATGCTTCAGATACCTACTTACAGTAGACACAGGCGCTA---CCCCTTGTCTCAGGTTTC 5085  
|||||  
5117 CACACATCGTACAGATCTTACTTACTCTCATGTACAGAAGAGAAATGGCATGGATCAAGAGA 5176  
|||||  
5086 AAGAAGACAGTGAAGAGGCTCATTCGGCATCTTCTTCTTTCGACAGAAATTTCTCAATATA 5145  
|||||  
5177 ATCCCACTGGGGAAGAAGAACTGGCCATTTCTCTTTGACAAAAACACTTCAACAGTATA 5236  
|||||  
5146 CAAGTGGCTCCATTTGGGTGTGATATCTCGTGGCTGTATTTTGGGTCTCACTGAAGAAATG 5205  
|||||  
5237 CTGGAAGTGCCTTGAAGAAGCCCTCGGCTGACACTGTGGAGGAGCGTCCGGGCGG 5296  
|||||  
5206 GCATCACCCCTTATCTCGCTGTGCTGATGTTTTCACATATTTTCTTGGGGTAACTC 5265  
|||||  
5297 CCATCATGCTTCTCCTCAAGTCTCTGCTGTTTCTTCTTCTTCACTATTTAAATGGAGTCCGGG 5356  
|||||  
5266 GCGCTGAGGAAGTCAATACCAATTTCTGCAAGGAGAGTACAGTGCATCTCTGTAGCTATC 5325  
|||||  
5357 CCCCTCCAGACCTTCA---AGTTTCTGGAACAAGCCATTTTGAACACTTATGTAACTACC 5413  
|||||  
5326 TATCTTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATGGGATACTGTAAAGCCCT 5385  
|||||  
5414 TTTCTTACCAACCACTCATCTGACCTTTTTCAGAAACACNGTGACATCATGACTCC 5473  
|||||  
5386 TGCTCCAGAGGTGTGTGCAGATCCTGCTTACTTAAACTGTTTGAAGCAAAAAACACCG 5445  
|||||  
5474 TGAATGAAGTTGGTGCAGAACAGTGAAGTTAAACGGTATCTAAATGGCGAGAGAGAG 5533  
|||||  
5446 TGCTCAGGTACCTAGAAAAGAAATAGTTTGTATAGAGTTCCTGTATGACTATAGTCC 5505  
|||||  
5534 CGATAAGCTTACCCAGAGAGAGTAAACAACTGATAGACTTCCAGAGGATATACAGACCC 5593  
|||||  
5506 TCCTGAATCAAGCTTCTCATTTTCAAGTGCACCGTCTGCAGATGATGAGCGAAAGCATC 5565  
|||||  
5594 TCAITTAACAGCATCCAAATTTCTGTGCCCCAAATCAGGTGGCGCAAGAGCAGGCTC 5653  
|||||  
5566 CTGCTCTCTGCTTTTCTGTGGGCTATACATATGTTCTCAGAAACATTTGCTGCCAGGAAA 5625  
|||||  
5654 CTACTCTGTGCTGTGTGGGAGTCTCCTCTGTCTCTCAGAGTTACTGCTGCCAAGCTG 5713  
|||||  
5626 TTGTGAACGGGAAGAGTTGAGCTTGCATTTTTCAGCCACTTCACTGTGGAGCGCGAG 5685  
|||||  
5714 AGCTGGAGGGTGAAGACGTCGAGGCTTGCAAGCAACACACTTACTCTCTCGCGCTCCGGG 5773  
|||||  
5686 TCTGCATTTTCTTAAATAATCAGAGAATGCCAGTGGTCTGTGTTGAAGTAAAGCCAGAG 5745  
|||||  
5774 CCGGCATCTTCTGAGGTGCGGGAATGTGAGTGTCTATTTTGTAGCTGGCAAAACCAAG 5833  
|||||  
5746 GCTGTGCTCTATCAGCTCCTTACTTTGGATGAATATGGAGAAACAGACCCCTGGCTGAAGA 5805  
|||||  
5834 GATGTTTTTTTATCTCTCTCTTACTTGACGACTATGGAGAGACCACAGGAGACTCAGAC 5893  
|||||  
5806 GGGGCAACCCCTTCATTTATCTCTGAGCGGTATCGGAGCTCCATTTGCTTGGTCCGAC 5865  
|||||  
5894 GAGGAATCTCTTTTATCATTTATGCCAAGAGCGGTTTTGGAAGATCAGAAAGCTCTGGCAGC 5953  
|||||  
5866 AACACTGCAATTATAGAAGAGATTGCTTAGGAGCAAGAGACTAATCAGATGTTTATTTGGAT 5925

Db 5954 AGCATAGTATCACAGAGAGATCGACACGCGCAGGAGGCTTAACAGACCCCTGGTCCGAA 6013  
|||||  
QY 5926 TCAACTGGCAGTTACTCTGTGAGCTC 5949  
|||||  
Db 6014 TTGACTGGCAGCATTTTATAATCGC 6037  
|||||  
RESULT 7  
US-08-982-956-2  
; Sequence 2, Application US/08982956  
; Patent No. 5861312  
; GENERAL INFORMATION:  
; APPLICANT: Varshavsky, Alexander  
; APPLICANT: Kwon, Yong Tae  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING MAMMALIAN UBRL  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kevin M. Farrell, P.C.  
; STREET: P.O. Box 999  
; CITY: York Harbor  
; STATE: ME  
; COUNTRY: US  
; ZIP: 03911  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/982,956  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farrell, Kevin M.  
; REGISTRATION NUMBER: 35,505  
; REFERENCE/DOCKET NUMBER: CIT-2001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (207) 363-0558  
; TELEFAX: (207) 363-0528  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1001 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..999  
US-08-982-956-2  
Query Match 15.8%; Score 996.2; DB 2; Length 1001;  
Best Local Similarity 99.7%; Pred. No. 5.2e-281;  
Matches 998; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2799 ATGATGCCCAATAAGTTCTTCTGTACTTCTCAGAGGTATGAACCTTGCAGGCTTTT 2858  
Db 1 ATGATGCCCAACAAGTTCTTGTACTTGTGTAATTCAGAGGTATGAACCTTGCAGGCTTTT 60  
QY 2859 AACAGACCATATCTACAAAAGACAGGATTTGATTAACAAATATAATACATAAGAA 2918  
Db 61 AACAGACCATATCTACAAAAGACAGGATTTGATTAACAAATATAATACATAAGAA 120  
QY 2919 GAAATGCTTCAGGTCCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGAAAT 2978  
Db 121 GAAATGCTTCAGGTCCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGAAAT 180  
QY 2979 GTGACCAAGAAGAGGTGACAAATCAATTCATCTTGTCTTGTGCAATTAAGAACCCATG 3038  
Db 181 GTGACCAAGAAGAGGTGACAAATCAATTCATCTTGTCTTGTGCAATTAAGAACCCATG 240  
QY 3039 CCACACAGTGCCTATTGCCAAAAATTTTACCTGAGAAATGAATAATGAATACTGGCTTAGAG 3098

Db 241 CCACAGTGCCTTGGCCAAATTTACCTGAGATGAAATTAATGAAGTGGCTTAGAG 300  
QY 3099 AATGTCTATAAACAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGAGTTTAT 3158  
Db 301 AATGTCTATAAACAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGAGTTTAT 360  
QY 3159 GAATCTAAAGATGATCATCTGAAGACCTTCAATATGTACTTTTATCATCTACTCCAAACC 3218  
Db 361 GAATCTAAAGATGATCATCTGAAGACCTTCAATATGTACTTTTATCATCTACTCCAAACC 420  
QY 3219 CAGCATAGCAGGCTGAACATATGAGAGAAAGAGAGAAACAAAGAAACAAAGATGAA 3278  
Db 421 CAGCATAGCAGGCTGAACATATGAGAGAAAGAGAGAAACAAAGAAACAAAGATGAA 480  
QY 3279 GCATTCGCCCCACACACCTCTCTGAATTCGCCCTGCTTTTCAGCAAAAGTGAATTAACCTT 3338  
Db 481 GCATTCGCCCCACACACCTCTCTGAATTCGCCCTGCTTTTCAGCAAAAGTGAATTAACCTT 540  
QY 3339 CTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAGGGGCAATAGACACA 3398  
Db 541 CTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAGGGGCAATAGACACA 600  
QY 3399 GATTCCTAACTTGTGACCGAGGGATGCTCCAAATGGCTTTTCATATCTGGCATTTGGGT 3458  
Db 601 GATTCCTAACTTGTGACCGAGGGATGCTCCAAATGGCTTTTCATATCTGGCATTTGGGT 660  
QY 3459 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTGAAAGAGAGAGTAAACATTTGACCTT 3518  
Db 661 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTGAAAGAGAGTAAACATTTGACCTT 720  
QY 3519 TATCTAAGGCTTCAAGATTTGGAGTTCAGCATGATATCAAAATGCTTTTGGAAAAA 3578  
Db 721 TATCTAAGGCTTCAAGATTTGGAGTTCAGCATGATATCAAAATGCTTTTGGAAAAA 780  
QY 3579 CTCAAGGATTCCTCCAGTTAGAGCCAGAGGACATGATTAACCTGATCTTACAGATG 3638  
Db 781 CTCAAGGATTCCTCCAGTTAGAGCCAGAGGACATGATTAACCTGATCTTACAGATG 840  
QY 3639 TTGTACAGATGAAGCGATTAAGAGAAATCTTGTAAATGTAGCAACACATCAGGA 3698  
Db 841 TTGTACAGATGAAGCGATTAAGAGAAATCTTGTAAATGTAGCAACACATCAGGA 900  
QY 3699 TCGGAATCTTAAAGATGATGAGATTTACTCATGATAAAGAAAAAGCAAGCAAAAAAGA 3758  
Db 901 TCGGAATCTTAAAGATGATGAGATTTACTCATGATAAAGAAAAAGCAAGCAAAAAAGA 960  
QY 3759 AAAGCTGAAGCTGTAGGCTAGCATGCCAGAGATCATGGC 3799  
Db 961 AAAGCTGAAGCTGTAGGCTTATCCGCCAGAGATCATGGC 1001

RESULT 8

US-09-228-317-2  
; Sequence 2, Application US/09228317  
; Patent No. 6159732  
; GENERAL INFORMATION:  
; APPLICANT: Varshavsky, Alexander  
; APPLICANT: Kwon, Yong Tae  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING MAMMALIAN UBR1  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kevin M. Farrell, P.C.  
; STREET: P.O. Box 999  
; CITY: York Harbor  
; STATE: ME  
; COUNTRY: US  
; ZIP: 03911  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/228,317  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farrell, Kevin M.  
; REGISTRATION NUMBER: 35,505  
; REFERENCE/DOCKET NUMBER: CIT-2001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (207) 363-0558  
; TELEFAX: (207) 363-0528  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1001 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..999  
; US-09-228-317-2  
  
Query Match 15.8%; Score 996.2; DB 3; Length 1001;  
Best Local Similarity 99.7%; Pred. No. 5.2e-281;  
Matches 998; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2799 ATGATCCCAATAAGTCTCTGTACTGTACTTCTCAGAGGTATGAACCTTGGCCAGGCTTTT 2858  
Db 1 ATGATCCCAACAAGTCTCTGTACTGTACTTCTCAGAGGTATGAACCTTGGCCAGGCTTTT 60  
  
QY 2859 AACAGACCATATCTACAAAAGACAGGATTTGATTAACAATATAATACATAAGAA 2918  
Db 61 AACAGACCATATCTACAAAAGACAGGATTTGATTAACAATATAATACATAAGAA 120  
  
QY 2919 GAATGCTTTCAGGTCCTCATCTATATTGCGGTGAGCGTTATGTACCTGGAGTGGGAAT 2978  
Db 121 GAATGCTTTCAGGTCCTCATCTATATTGCGGTGAGCGTTATGTACCTGGAGTGGGAAT 180  
  
QY 2979 GTGACCAAGAGAGGTGACATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCCATG 3038  
Db 181 GTGACCAAGAGAGGTGACATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCCATG 240  
  
QY 3039 CCACACAGTGCCTTGGCCAAATTTTACCTCAGAAATGAAAATAATGAACCTGGCTTAGAG 3098  
Db 241 CCACACAGTGCCTTGGCCAAATTTTACCTCAGAAATGAAAATAATGAACCTGGCTTAGAG 300  
  
QY 3099 AATGCTATAACAAGTGGCCACATTTTAAGAAACAGGTGTATCAGGCCATGGAGTTTAT 3158  
Db 301 AATGCTATAACAAGTGGCCACATTTTAAGAAACAGGTGTATCAGGCCATGGAGTTTAT 360  
  
QY 3159 GAATCTAAAGATGATCATCTGAAGACCTTCAATATGTACTTTTATCATTTCTCCAAACC 3218  
Db 361 GAATCTAAAGATGATCATCTGAAGACCTTCAATATGTACTTTTATCATTTCTCCAAACC 420  
  
QY 3219 CAGCATAGCAGGCTGAACATATGAGAGAAAGAGAGAAACAAAGAAACAAAGATGAA 3278  
Db 421 CAGCATAGCAGGCTGAACATATGAGAGAAAGAGAGAAACAAAGAAACAAAGATGAA 480  
  
QY 3279 GCATTCGCCCCACACACCTCTCTGAATTCGCCCTGCTTTTCAGCAAAAGTGAATTAACCTT 3338  
Db 481 GCATTCGCCCCACACACCTCTCTGAATTCGCCCTGCTTTTCAGCAAAAGTGAATTAACCTT 540  
  
QY 3339 CTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGGGGCAATAGACACA 3398  
Db 541 CTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGGGGCAATAGACACA 600  
  
QY 3399 GATTCCTAACTTGTGACCGAGGGATGCTCCAAATGGCTTTTCATATTTCTGGCATTTGGGT 3458  
Db 601 GATTCCTAACTTGTGACCGAGGGATGCTCCAAATGGCTTTTCATATTTCTGGCATTTGGGT 660  
  
QY 3459 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTGAAAGAGAGTAAACATTTGACCTT 3518  
Db 661 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTGAAAGAGAGTAAACATTTGACCTT 720

661	Db	TTACTAGAGAGAACCAACGCTTTCAAAAGCTCCTGAAGAGAAGTAACTTTGACTTT	720
3519	Qy	TATCATAAAGGCTTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAA	3578
721	Db	TATCATAAAGGCTTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAA	780
3579	Qy	CTCAAGAGGAATCCCCAGTTTAGAAGGCCAAGAGCATGATACGTGGATCTTCAGATG	3638
781	Db	CTCAAGAGGAATTTCCCGATTGAAAGGCCAGAGGACATGATTAACGTGGATCTTCAGATG	840
3639	Qy	TTTGACACACTGAAGCGATTATGAGAGAAAAATCTTGTTTAATTTGAGCAACCATCAGGA	3698
841	Db	TTTGACACACTGAAGCGATTATGAGAGAAAAATCTTGTTTAATTTGAGCAACCATCAGGA	900
3699	Qy	TCCGGAATCTAATTAAGAAATGATGAGTTACTCATGATATAAGCAAAAAAGCAGAACCGAAAAAGA	3758
901	Db	TCCGAATCTAATTAAGAAATGATGAGTTACTCATGATATAAGCAAAAAAGCAGAACCGAAAAAGA	960
3759	Qy	AAAGCTGAAGCTGCTAGGCTACATCGCCAGAAGATCATGCG	3799
961	Db	AAAGCTGAAGCTGCTAGGCTTCATCGCCAGAAGATCATGCG	1001

## RESULT 9

```

US-09-270-767-14761
; Sequence 14761, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and
; FILE REFERENCE: File Reference: 7326-0
; CURRENT APPLICATION NUMBER: US/09/270,
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14761
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14761

```

Query Match	1.8%	Score 113.4	DB 4	Length 670
Best Local Similarity	53.4%	Pred. No. 3e-22		
Matches 287	Conservative 0	Mismatches 241	Indels 9	Gaps 2
Qy	2943	ATTGTGGGTGAGCGCTTATGTACCTGGAGTGGGAAATGTGACCAAGAGAGGTCACAAATG	3002	
Db	2	ATCATTTGGAGAACTTTGGATGCCCTGGCGTTTCGATGTGTGACCGNAGAGGATCGCTGGCG	61	
Qy	3003	AGAGAAATCAATTCATTCTGTTTGGATTTGAACCCATGCCACACAGTGCATTTGCCAAAAAT	3062	
Db	62	AAGGAGATCATCCAGCTGTTGTGTATCAAGCCATACTCACACTCGGAGTTGAGTCGCGCC	121	
Qy	3063	TTACCTG-----AGATGAATAATAATGAACCTGGCTTACGAGATGTTCATAAACAAAGTG	3116	
Db	122	TTGCAGATGGCAATAGCGGAATAGCGCAACGTTTTTCGAGGAAGTTATCAACACGGTT	181	
Qy	3117	GCCACATTTAAGAAACCAGGTGTATCAG---GCCATGGAGTTTATGAACTAAAGAGTAA	3173	
Db	182	GCGTGTTCAGAAACCCTGTTGGAGCCGACAGCAAGGGGTGTACGAACTAAAGGAACAT	241	
Qy	3174	TCACAGAAAGACTTCAATATGTACTTTTATCATTACTCCAAAAACCCAGCATAGCAAGGCT	3233	
Db	242	TTGCTTAAGGAGTTTCAACATGTACTTCTATCACTACCAAGAGAGGCAAGTCCNAGGCC	301	
Qy	3234	GMACATATGCAGAGAAAGAGAGAAAAACAAGAAACAAGCATTTGCCCGCCACCA	3293	
Db	302	GAGGACTGCACCGGAGCGCCGAAAGGCCAGAGACGCTGTGTCTGCTGCCCGCCACCA	361	
Qy	3294	CCACTCTGAAATTCCTCCCTGCTTTTCAGCAAAAGTGATTAACCTTCTCAACTGTGATATC	3353	
Db	362	ATGTTGCAGAAACTTAAACAACAGCCTTTCACACCATGGCCCAACACTCTTCGAGTGCCTGTG	421	

Qy	3354	ATGATGTACATTTCTCAGGACCGTATTTTGAGCGGGCATATACACACAGATTCTTAATCTTGTGG	3413
Db	422	TTCTTCAACATCTGCAGCTTGATTATGGAAGGGCTTTAAATGCTTACAGTCGCAGCTTT	481
Qy	3414	ACCGAAGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTGGGTTTTACTAGAAAGAG	3470
Db	482	ACCGAATCGCATCTTCAAAGGTCCTTCACCTGCTGGGTTATGCCATTCAGGAGGAG	538

RESULT 10

US-09-513-999C-936/c

; Sequence 936, Application US/09513999C

; Patent No. 6783961

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6783961

**RESULT 10**

US-09-513-999C-936/c  
 ; Sequence 936, Application US/0951399C  
 ; Patent No. 6783961  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Duclert, A.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 ; Patent No. 6783961

```

FILE REFERENCE: 59.US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent_pm
SEQ ID NO 936
LENGTH: 366
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 37..366
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8
OTHER INFORMATION: n=a, g, c o r t
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9
OTHER INFORMATION: n=a, g, c o r t
FEATURE:
NAME/KEY: misc_feature
LOCATION: 317
OTHER INFORMATION: n=a, g, c o r t
FEATURE:
NAME/KEY: misc_feature
LOCATION: 318
OTHER INFORMATION: k=g o r t
FEATURE:
NAME/KEY: UNSURE
LOCATION: 94
OTHER INFORMATION: Xaa=His or Leu or Pro
US-09-513-999C-936

```

	Query Match	1.7%;	Score 105.6;	DB 4;	Length 366;
	Best Local Similarity	82.7%;	Pred. No. 3.9e-20;		
	Matches 134;	Conservative 0;	Mismatches 24;	Indels 4;	Gaps 1;
Qy	123	AAAGGCTTATCATCTGTCCTTCCACTTATCCAAAGCTGCTATGCGGCACCTGCCGTGCGG	182		
Db	173	ACAGGCTTATCATCTATCTTCCACTTATCCAAAGCGCGGTACGGCCACTTCCTGTGCGG	114		
Qy	183	CACCTGGAACACACGGCCAGCGCCCACTACTGCTCCACTACACATGCTTCTCCCAACCCCTGA	242		
Db	113	CAGTTGGAAGCACACACGAGCCCACTACAGGCTCCGCTACACCGCTTCTCCCGCCCTGT	54		
Qy	243	TCAGCTGCTTGCTGCTGCCACTTATCCGCTTTCGCTGTTC	284		
b	53	ACAGC-----TGCGCGCCCACTTGTCCGCTCTGCGCTCTCC	16		

## RESULT 11

US-09-513-999C-938/c  
; Sequence 938, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 938  
; LENGTH: 456  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 37..456  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 8  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 9  
; OTHER INFORMATION: n=a, g, c or t  
; US-09-513-999C-938

Query Match 1.7%; Score 105.6; DB 4; Length 456;  
Best Local Similarity 82.7%; Pred. No. 4.5e-20;  
Matches 134; Conservative 0; Mismatches 24; Indels 4; Gaps 1;

QY 123 AAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGCTATGCGCACTGCTGTGCGG 182  
DB 173 ACAGGCTTATCATCTATCTTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCGG 114  
QY 183 CACCTGGAAACACCGCCGACCTACTGCTTCCACTACCACTGGTCTCCACCCTGA 242  
DB 113 CAGTTGGAAGCACCACCGCCGCTACAGCCTCGCTACCACTGCTTCTCCGCCCTGT 54  
QY 243 TCAGCTGCTTGTCTGCTGCTATCTTCCGCTTCTGCTGCTTC 284  
DB 53 ACAGC-----TGGCGCGGCACTTGTGCGGCTCTGCTCTCC 16

RESULT 12  
US-09-513-999C-937/c  
; Sequence 937, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 937  
; LENGTH: 605  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: 177..605  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 8  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 9  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 49  
; OTHER INFORMATION: s=g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 199  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 8  
; OTHER INFORMATION: Xaa=Ala or Asp or Gly or Val  
; US-09-513-999C-937

Query Match 1.6%; Score 103.2; DB 4; Length 605;  
Best Local Similarity 81.5%; Pred. No. 2.8e-19;  
Matches 132; Conservative 1; Mismatches 24; Indels 5; Gaps 1;

QY 123 AAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGCTATGCGCACTGCTGTGCGG 182  
DB 172 ACAGGCTTATCATCTATCTTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCGG 113  
QY 183 CACCTGGAAACACCGCCGACCTACTGCTTCCACTACCACTGGTCTCCACCCTGA 242  
DB 112 CAGTTGGAAGCACCACCGCCGCTACAGCCTCGCTACCACTGCTTCTCCGCCCTGT 53  
QY 243 TCAGCTGCTTGTCTGCTGCTATCTTCCGCTTCTGCTGCTTC 284  
DB 52 ACAGC-----TGGCGCGGCACTTGTGCGGCTCTGCTCTCC 16

RESULT 13  
US-09-513-999C-14808/c  
; Sequence 14808, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14808  
; LENGTH: 602  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 8  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 9  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 266  
; OTHER INFORMATION: k=g or t

; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 553  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 554  
; OTHER INFORMATION: k=g or t  
US-09-513-999C-14809

Query Match 1.6%; Score 100.8; DB 4; Length 602;  
Best Local Similarity 86.7%; Pred. No. 1.4e-18;  
Matches 111; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Qy 123 AAAGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGGCCACTGCTGTGCG 182  
Db 259 ACAGGCTTATCATCTATCTTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCCG 200  
Qy 183 CACCTGGAACCAACCGCCAGCCCACTACTGCTCCCACTACCACTGGTTCTCCCACTGA 242  
Db 199 CAGTTGGAAGCAACCAACCGCCCACTACAGCTCCGCTACCAACCGCTTCTCCGCGCTGT 140  
Qy 243 TCAGCTGC 250  
Db 139 ACAGCTGC 132

RESULT 14  
US-09-513-999C-14809/c  
; Sequence 14809, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14809  
; LENGTH: 692  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 8  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 9  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 266  
; OTHER INFORMATION: k=g or t  
US-09-513-999C-14809

Query Match 1.6%; Score 100.8; DB 4; Length 692;  
Best Local Similarity 86.7%; Pred. No. 1.5e-18;  
Matches 111; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Qy 123 AAAGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGGCCACTGCTGTGCG 182  
Db 259 ACAGGCTTATCATCTATCTTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCCG 200  
Qy 183 CACCTGGAACCAACCGCCAGCCCACTACTGCTCCCACTACCACTGGTTCTCCCACTGA 242  
Db 199 CAGTTGGAAGCAACCAACCGCCCACTACAGCTCCGCTACCAACCGCTTCTCCGCGCTGT 140

Qy 243 TCAGCTGC 250  
Db 139 ACAGCTGC 132  
RESULT 15  
US-09-513-999C-14947/c  
; Sequence 14947, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14947  
; LENGTH: 842  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 8  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 9  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 266  
; OTHER INFORMATION: k=g or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 436  
; OTHER INFORMATION: n=a, g, c or t  
US-09-513-999C-14947

Query Match 1.6%; Score 100.8; DB 4; Length 842;  
Best Local Similarity 86.7%; Pred. No. 1.8e-18;  
Matches 111; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Qy 123 AAAGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGGCCACTGCTGTGCG 182  
Db 259 ACAGGCTTATCATCTATCTTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCCG 200  
Qy 183 CACCTGGAACCAACCGCCAGCCCACTACTGCTCCCACTACCACTGGTTCTCCCACTGA 242  
Db 199 CAGTTGGAAGCAACCAACCGCCCACTACAGCTCCGCTACCAACCGCTTCTCCGCGCTGT 140  
Qy 243 TCAGCTGC 250  
Db 139 ACAGCTGC 132

Search completed: July 9, 2005, 08:02:28  
Job time : 1015 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2005, 02:53:40 ; Search time 3333 Seconds  
(without alignments)  
11882.270 Million cell updates/sec

Title: US-10-758-672A-1

Perfect score: 6308  
Sequence: 1 gccaaagattcgacgagg.....aattttgtattgtgtttt 6308

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/FCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6308	100.0	6308	19	US-10-758-672A-1
2	6308	100.0	6308	21	US-10-758-672A-1
3	5592	88.6	7742	16	US-10-287-218-22
4	5592	88.6	7742	16	US-10-474-291-22
5	5466	86.7	5466	20	US-10-357-819-9
6	5136	81.4	5205	21	US-10-758-672A-18
7	5136	81.4	5205	21	US-10-758-672A-18

Sequence 1, Appli  
Sequence 1, Appli  
Sequence 22, Appli  
Sequence 22, Appli  
Sequence 9, Appli  
Sequence 18, Appli  
Sequence 18, Appli

8	3055.8	48.4	3059	17	US-10-094-749-916	Sequence 916, Appl	
9	1381.8	21.9	6300	19	US-10-758-672A-3	Sequence 3, Appli	
10	1381.8	21.9	6300	21	US-10-758-672A-3	Sequence 3, Appli	
11	1233.6	19.6	6089	19	US-10-758-672A-5	Sequence 5, Appli	
12	1233.6	19.6	6089	21	US-10-758-672A-5	Sequence 5, Appli	
13	981	15.6	6158	19	US-10-755-889-634	Sequence 634, App	
14	813.6	12.9	3327	9	US-09-529-063-56	Sequence 56, Appl	
15	813.6	12.9	3327	16	US-10-414-378-56	Sequence 56, Appl	
16	813.6	12.9	3502	9	US-09-529-063-57	Sequence 57, Appl	
17	813.6	12.9	3502	16	US-10-414-378-57	Sequence 57, Appl	
18	438.4	6.9	505	10	US-09-918-995-36721	Sequence 36721, A	
19	374.8	5.9	1772	17	US-10-264-237-1139	Sequence 1139, Ap	
20	250.6	4.0	972	9	US-09-822-849A-532	Sequence 532, App	
C	21	219.8	3.5	4573	13	US-10-071-766-7	Sequence 7, Appli
22	212.2	3.4	578	13	US-10-027-632-260533	Sequence 260533,	
23	212.2	3.4	578	17	US-10-027-632-260533	Sequence 260533,	
24	160.2	2.5	3411	17	US-10-104-047-183	Sequence 183, App	
25	155.6	2.5	469	17	US-10-242-535A-26847	Sequence 26847, A	
26	155.6	2.5	469	18	US-10-085-783A-26847	Sequence 26847, A	
27	142	2.3	643	17	US-10-264-237-108	Sequence 108, App	
C	28	136.6	2.2	550	16	US-10-029-386-12897	Sequence 12897, A
29	117	1.9	121	16	US-10-029-386-12897	Sequence 12897, A	
30	116.2	1.8	364	17	US-10-242-535A-21996	Sequence 21996, A	
31	116.2	1.8	364	18	US-10-085-783A-21996	Sequence 21996, A	
C	32	105.6	1.7	1459	17	US-10-264-049-746	Sequence 746, App
33	105.6	1.7	1451	9	US-09-925-301-490	Sequence 490, Appl	
C	34	105.2	1.7	1692	20	US-10-335-053-82	Sequence 82, Appl
C	35	104.8	1.7	507	20	US-10-357-930-59123	Sequence 59123, A
C	36	100.8	1.6	697	9	US-09-823-245A-172	Sequence 172, App
C	37	100.8	1.6	1370	19	US-10-775-169-131	Sequence 131, App
C	38	100.8	1.6	1370	20	US-10-684-422-154	Sequence 154, App
C	39	100.8	1.6	1370	21	US-10-956-157-1949	Sequence 1949, Ap
C	40	100.8	1.6	1370	21	US-10-956-157-1949	Sequence 1949, Ap
C	41	100.6	1.6	1382	14	US-10-016-349A-118	Sequence 118, App
C	42	96	1.5	436	21	US-10-278-698-446	Sequence 446, App
C	43	96	1.5	436	21	US-10-278-698-446	Sequence 446, App
44	95.2	1.5	516	21	US-10-278-698-457	Sequence 457, App	
45	95.2	1.5	516	21	US-10-278-698-972	Sequence 972, App	

#### ALIGNMENTS

RESULT 1  
US-10-758-672A-1  
; Sequence 1, Application US/10758672A  
; Publication No. US20040185037A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966B  
; CURRENT APPLICATION NUMBER: US/10758.672A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 1  
; LENGTH: 6308  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (696)..(5942)  
US-10-758-672A-1

Query Match 100.0%; Score 6308; DB 19; Length 6308;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCCAAGAATTCGGCCAGGGGGAAGCTGAGCCAGAACCAATTAATTCTTGTACTCT 60

|||||  
1 GCCAAGAAATTCGGACAGAGGGGAAAGCTGAGCCAGGAACCAAAATTACTTGGCTTACCT 60  
QY  
61 CATTGTGTAAGACAAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAGAACTTACCT 120  
Db  
61 CATTGTGTAAGACAAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAGAACTTACCT 120  
QY  
121 CCAAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGCTATGGCCACTGCTGTGC 180  
Db  
121 CCAAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGCTATGGCCACTGCTGTGC 180  
QY  
181 CGCACTTGGAAACCAACCGGACGCCCACTACTGCTTCCACTACCACTGGTTCTCCACCCT 240  
Db  
181 CGCACTTGGAAACCAACCGGACGCCCACTACTGCTTCCACTACCACTGGTTCTCCACCCT 240  
QY  
241 GATCAGCTGCTTGTGCTGCCATCTTATCCGCTTCTGCTGTCTGAGTAATGTATACA 300  
Db  
241 GATCAGCTGCTTGTGCTGCCATCTTATCCGCTTCTGCTGTCTGAGTAATGTATACA 300  
QY  
301 CACCTTGGAAACCAACCTTACTTCTGTGCTATGAATTTGACTACTCTAGCTGGATC 360  
Db  
301 CACCTTGGAAACCAACCTTACTTCTGTGCTATGAATTTGACTACTCTAGCTGGATC 360  
QY  
361 CGAGCTTTTGTACACATGTGCAAGTGCACCGGGGTAGAATCCTTAAAAATAGAAGAT 420  
Db  
361 CGAGCTTTTGTACACATGTGCAAGTGCACCGGGGTAGNA TCCTTAAAAATAGAAGAT 420  
QY  
421 GTATGCAACAGTTCCTCCAGCAACCAACCGAGATATACAA CCAATTCAGCTACCAAGACTAC 480  
Db  
421 GTATGCAACAGTTCCTCCAGCAACCAACCGAGATATACAA CCAATTCAGCTACCAAGACTAC 480  
QY  
481 GCCTGATAAATTAGAGGGGAAAAAATAATCTCCAGTCCCTTCAAGTCCCTGCTGACGCTTGT 540  
Db  
481 GCCTGATAAATTAGAGGGGAAAAAATAATCTCCAGTCCCTTCAAGTCCCTGCTGACGCTTGT 540  
QY  
541 TCCGGGAAGCGGCGGGAAGCACTCTCGAGTCTGCGTCAAAACCCGACTTTCAGGGGCGG 600  
Db  
541 TCCGGGAAGCGGCGGGAAGCACTCTCGAGTCTGCGTCAAAACCCGACTTTCAGGGGCGG 600  
QY  
601 TCGTAAAAAGTGTGCTCCTGTCTCCGACGGGCCACAGGTTTCGCTTGGCCG 660  
Db  
601 TCGTAAAAAGTGTGCTCCTGTCTCCGACGGGCCACAGGTTTCGCTTGGCCG 660  
QY  
661 GGGGTCCGCACTGACGGGTCAAGTTCCCTCAAGATGGCGACGAGGCGTGGAGTA 720  
Db  
661 GGGGTCCGCACTGACGGGTCAAGTTCCCTCAAGATGGCGACGAGGCGTGGAGTA 720  
QY  
721 CTGAGAGGATGGAATCAGCGCGGAGTTACCCGACAGCCCTCAGCGTCTGGCATCTTGGT 780  
Db  
721 CTGAGAGGATGGAATCAGCGCGGAGTTACCCGACAGCCCTCAGCGTCTGGCATCTTGGT 780  
QY  
781 GGGATCAGCAAGTTGAATTTTATCTGCTTCTTGCATCAATTTGGCACAATTTGGTGGCAG 840  
Db  
781 GGGATCAGCAAGTTGAATTTTATCTGCTTCTTGCATCAATTTGGCACAATTTGGTGGCAG 840  
QY  
841 AAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTACAAATGT 900  
Db  
841 AAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTACAAATGT 900  
QY  
901 CAATATTCTCACTCGAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960  
Db  
901 CAATATTCTCACTCGAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960  
QY  
961 AATTGAAGCAGAGTGGAGCAATTCAGCTTTGTGGAGGGTTTCAAAAGTGGAGAGACAA 1020  
Db  
961 AATTGAAGCAGAGTGGAGCAATTCAGCTTTGTGGAGGGTTTCAAAAGTGGAGAGACAA 1020  
QY  
1021 CCTATTCTTGAGGATTTGCAATTTGATCCAACTGTGTA CTCTGTATGGACTGCTTCC 1080  
Db  
1021 CCTATTCTTGAGGATTTGCAATTTGATCCAACTGTGTA CTCTGTATGGACTGCTTCC 1080  
QY  
1081 AGGACAGATGTTCAATAAATCATCGTTACAAGATGCATCTTCTACTGGAGGAGGTTCT 1140  
|||||

Db  
1081 AGGACAGTGTTCATAAAAAATCATCTTTACAAGATGCATCTTCTACTGGAGGAGGTTCT 1140  
QY  
1141 GTGACTCTGGAGACACAGAGSCATGGAAAACCTGCCCTTTTGTGTAATCATCAACCTG 1200  
Db  
1141 GTGACTCTGGAGACACAGAGSCATGGAAAACCTGCCCTTTTGTGTAATCATCAACCTG 1200  
QY  
1201 GAAGCAGAGGTACTATATAAAGAGAAATTCAGCGTGTCCGTTGAAATGAAGAGGTAA TTGTCC 1260  
Db  
1201 GAAGCAGAGGTACTATATAAAGAGAAATTCAGCGTGTCCGTTGAAATGAAGAGGTAA TTGTCC 1260  
QY  
1261 AAGCAGAGAAAATATTTCTTCTTCACTGATAAATATGTCGTAGAAAATGACTATATGGGAAG 1320  
Db  
1261 AAGCAGAGAAAATATTTCTTCTTCACTGATAAATATGTCGTAGAAAATGACTATATGGGAAG 1320  
QY  
1321 AGGAAAAGAACTCCCTCCCTGAACTCCAGATAAGGGAGAAAAATGAAGATFAC TATTGTG 1380  
Db  
1321 AGGAAAAGAACTCCCTCCCTGAACTCCAGATAAGGGAGAAAAATGAAGATFAC TATTGTG 1380  
QY  
1381 TCCTTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGGCTCA CAAAGAGCTC 1440  
Db  
1381 TCCTTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGGCTCA CAAAGAGCTC 1440  
QY  
1441 TTGACTGTGAGCTCGCAGAGGCCCAAGTTGCATACCACTGCCATTTGCA CAAAGAGGGTCTC 1500  
Db  
1441 TTGACTGTGAGCTCGCAGAGGCCCAAGTTGCATACCACTGCCATTTGCA CAAAGAGGGTCTC 1500  
QY  
1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGAGATATAAAGAGCTC 1560  
Db  
1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGAGATATAAAGAGCTC 1560  
QY  
1561 ATTCAAGAAAATGTCTCTCAACATCCACTTCATGTAGAGTATTACACTCAGAGATTATGG 1620  
Db  
1561 ATTCAAGAAAATGTCTCTCAACATCCACTTCATGTAGAGTATTACACTCAGAGATTATGG 1620  
QY  
1621 CTCAATCAAGAAATTTGCTTTGCTGCTTGGTCTCTCGATGAA CAAAATTTATAGCTATTCAA 1680  
Db  
1621 CTCAATCAAGAAATTTGCTTTGCTGCTTGGTCTCTCGATGAA CAAAATTTATAGCTATTCAA 1680  
QY  
1681 GTGACTTTAGCAGACTTTTGGCCAGCATGCCCTTAGAGAGAACCTGACTCGGAGATC 1740  
Db  
1681 GTGACTTTAGCAGACTTTTGGCCAGCATGCCCTTAGAGAGAACCTGACTCGGAGATC 1740  
QY  
1741 CCTGTCTCATAAAGCAGGTTAATGCTTTGGGATGCAAAAGCTTTATAAAGGTGCCGTAAAG 1800  
Db  
1741 CCTGTCTCATAAAGCAGGTTAATGCTTTGGGATGCAAAAGCTTTATAAAGGTGCCGTAAAG 1800  
QY  
1801 TCCTTCATGAATTTGATCTTCAAGAGTTTATTTATGAGATGGAATATATCATGATGATCATG 1860  
Db  
1801 TCCTTCATGAATTTGATCTTCAAGAGTTTATTTATGAGATGGAATATATCATGATGATCATG 1860  
QY  
1861 CTATGGAATTTGTGAAGTATTATAAAACCTGCAAGAAAGATATATCATGATGATCATG 1920  
Db  
1861 CTATGGAATTTGTGAAGTATTATAAAACCTGCAAGAAAGATATATCATGATGATCATG 1920  
QY  
1921 ACAGAGTATCTCTATAACCTGCACTTTTTCAGTTCAGATGTTTACTGTTCTCTACTCTGGCTC 1980  
Db  
1921 ACAGAGTATCTCTATAACCTGCACTTTTTCAGTTCAGATGTTTACTGTTCTCTACTCTGGCTC 1980  
QY  
1981 GACATCTTATTTGAAGAGCAGAAATGTTATCTCTGCTCA TTACTGAAACCTCTGCTAGAGTTT 2040  
Db  
1981 GACATCTTATTTGAAGAGCAGAAATGTTATCTCTGCTCA TTACTGAAACCTCTGCTAGAGTTT 2040  
QY  
2041 TACCTGAGTACTCTGACAGGAACTAAATTTCACTTCCAGGGTTATAGCCAGGCAAAAT 2100  
Db  
2041 TACCTGAGTACTCTGACAGGAACTAAATTTCACTTCCAGGGTTATAGCCAGGCAAAAT 2100  
QY  
2101 TGGGAAGAGTATATGCACTAAATGTGACCTTAAAGTATATCTGATCAGCAAAACCCCAA 2160  
Db  
2101 TGGGAAGAGTATATGCACTAAATGTGACCTTAAAGTATATCTGATCAGCAAAACCCCAA 2160  
QY  
2161 TATGACAGAAAGATTAAGAAATGCAAGTTCCCTTGAAGGTTTTCGATCTTTTGTGAAGTTC 2220  
Db  
2161 TATGACAGAAAGATTAAGAAATGCAAGTTCCCTTGAAGGTTTTCGATCTTTTGTGAAGTTC 2220

QY 2221 TTACTGTATGACGGGATGGAAGAAATCCGAGACAGGTTGGGCAACACATTGAAAGTG 2280  
DB 2221 TTACTGTATGACGGGATGGAAGAAATCCGAGACAGGTTGGGCAACACATTGAAAGTG 2280  
QY 2281 ATCTCTGATTGGGAGGCTGCCATTGCTATACAGATGCAATTTGAAGAAATATTTTACTCATGT 2340  
DB 2281 ATCTCTGATTGGGAGGCTGCCATTGCTATACAGATGCAATTTGAAGAAATATTTTACTCATGT 2340  
QY 2341 TCCAGAGTGGTGTGCTTGTGTGATGAAGAACTCTTACTTGTGCTTATTAAGAAATGTCA 2400  
DB 2341 TCCAGAGTGGTGTGCTTGTGTGATGAAGAACTCTTACTTGTGCTTATTAAGAAATGTCA 2400  
QY 2401 AGCTGTGATGAGTGCAGTACAGTCTTATCTAGTAGCAAGACAGTAGTACAAATCGT 2460  
DB 2401 AGCTGTGATGAGTGCAGTACAGTCTTATCTAGTAGCAAGACAGTAGTACAAATCGT 2460  
QY 2461 GTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTGAGGATCTTGTGAACATACATC 2520  
DB 2461 GTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTGAGGATCTTGTGAACATACATC 2520  
QY 2521 TGCCACTCTCTAGGACCCCTTGTGCTCTTCACTAGTCTTAAAGCAGGCTGGTGTCTT 2580  
DB 2521 TGCCACTCTCTAGGACCCCTTGTGCTCTTCACTAGTCTTAAAGCAGGCTGGTGTCTT 2580  
QY 2581 CAGACTGATGAATTTGTGCTCTTGTAGGACTTTCAAGTAGAGTACTAGTGGAAATATC 2640  
DB 2581 CAGACTGATGAATTTGTGCTCTTGTAGGACTTTCAAGTAGAGTACTAGTGGAAATATC 2640  
QY 2641 CTTTACGTTGCTGCTGTGTTGGTCCAGTGTGCTGAGATGCGGAGAAATGGAC 2700  
DB 2641 CTTTACGTTGCTGCTGTGTTGGTCCAGTGTGCTGAGATGCGGAGAAATGGAC 2700  
QY 2701 TGTCTCTTATTAGCCAGGTGTTTATTATACCAAGATGTTAAAGTGCAGAGAAATGTATG 2760  
DB 2701 TGTCTCTTATTAGCCAGGTGTTTATTATACCAAGATGTTAAAGTGCAGAGAAATGTATG 2760  
QY 2761 ATAAAGATATCATATGCTTCAGATGTTGCTGATCTTTAATGATCCCAATTAAGTCTTGT 2820  
DB 2761 ATAAAGATATCATATGCTTCAGATGTTGCTGATCTTTAATGATCCCAATTAAGTCTTGT 2820  
QY 2821 TACTGGTACTTCAGAGGTATGAACCTTGGGAGGCTTTTAAACAGACCATATCTACAAAG 2880  
DB 2821 TACTGGTACTTCAGAGGTATGAACCTTGGGAGGCTTTTAAACAGACCATATCTACAAAG 2880  
QY 2881 ACCGAGTTTGAATTAACAAATAATACTAATAAGAAATGCTTCAGGTCCTCATCT 2940  
DB 2881 ACCGAGTTTGAATTAACAAATAATACTAATAAGAAATGCTTCAGGTCCTCATCT 2940  
QY 2941 ATATTGGGTGAGGTTATGTAATGTAAGTGGAAATGTGACCAAGAGAGGTCA 3000  
DB 2941 ATATTGGGTGAGGTTATGTAATGTAAGTGGAAATGTGACCAAGAGAGGTCA 3000  
QY 3001 TGAGAGAAATCAATTCATCTGCTTTCATGTAAGCCATGCCACAGTGGCCATTGCCAAA 3060  
DB 3001 TGAGAGAAATCAATTCATCTGCTTTCATGTAAGCCATGCCACAGTGGCCATTGCCAAA 3060  
QY 3061 ATTTACCTGAGAAATCAATTAAGTGAAGTGGTGTGAGAAATGTCAATAAGAGTGGCA 3120  
DB 3061 ATTTACCTGAGAAATCAATTAAGTGAAGTGGTGTGAGAAATGTCAATAAGAGTGGCA 3120  
QY 3121 CATTTAAGAAACAGGTTATCAGGCCATGAGGTTTATGAATTAAGAAATGAATCACTGA 3180  
DB 3121 CATTTAAGAAACAGGTTATCAGGCCATGAGGTTTATGAATTAAGAAATGAATCACTGA 3180  
QY 3181 AAGACTCAATATGTAATTTTATCAATCTCAAAACCCAGCATAGCAGGCTGAACATA 3240  
DB 3181 AAGACTCAATATGTAATTTTATCAATCTCAAAACCCAGCATAGCAGGCTGAACATA 3240  
QY 3241 TGCAGAGAAAGGAGAGAAACAGAGAAACAGAGTGAAGCAATGCGCCACACACCTC 3300  
DB 3241 TGCAGAGAAAGGAGAGAAACAGAGAAACAGAGTGAAGCAATGCGCCACACACCTC 3300

QY 3301 CTGAATTTCTGCCCTGCTTTTCAGCAAGTGAATTAACCTTCTCAACTGTGATATCATGATGT 3360  
DB 3301 CTGAATTTCTGCCCTGCTTTTCAGCAAGTGAATTAACCTTCTCAACTGTGATATCATGATGT 3360  
QY 3361 ACATTTCTCAGACCCGCTATTTGAGCGGGCAATAGACACAGATTTCTAACTTGTGGACCGAAG 3420  
DB 3361 ACATTTCTCAGACCCGCTATTTGAGCGGGCAATAGACACAGATTTCTAACTTGTGGACCGAAG 3420  
QY 3421 GGATGCTCCAAATGGCTTTTCATATCTGCGCATTTGGGTTTACTAGAGAGAAACACAGC 3480  
DB 3421 GGATGCTCCAAATGGCTTTTCATATCTGCGCATTTGGGTTTACTAGAGAGAAACACAGC 3480  
QY 3481 TTCAAAAAGCTCTCGAAGNAGAAATCAATTTGACTTTTATCATAGGCTTCAAGATTTGG 3540  
DB 3481 TTCAAAAAGCTCTCGAAGNAGAAATCAATTTGACTTTTATCATAGGCTTCAAGATTTGG 3540  
QY 3541 GAAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAACTCAAGGAAATTCGCCAGTTAG 3600  
DB 3541 GAAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAACTCAAGGAAATTCGCCAGTTAG 3600  
QY 3601 AAGGCCAGAGGACATGATAACGTTGATACTTCAGATGTTTGACACAGTGAAGCGATTA 3660  
DB 3601 AAGGCCAGAGGACATGATAACGTTGATACTTCAGATGTTTGACACAGTGAAGCGATTA 3660  
QY 3661 GAGAAAAATCTTGTTTAATTTAGCAACCATCAGGATCGGAATCTATTAAAGATGATG 3720  
DB 3661 GAGAAAAATCTTGTTTAATTTAGCAACCATCAGGATCGGAATCTATTAAAGATGATG 3720  
QY 3721 AGATTACTCATGATAAAGAAAGCAGAAACGAAAAAGAAAGCTGAAGCTGTAGGCTAC 3780  
DB 3721 AGATTACTCATGATAAAGAAAGCAGAAACGAAAAAGAAAGCTGTAGGCTAC 3780  
QY 3781 ATGCGCAGAGATCATGCTCAGATGCTGCTTACAGAAAAATTTCAATTTGAATTCATA 3840  
DB 3781 ATGCGCAGAGATCATGCTCAGATGCTGCTTACAGAAAAATTTCAATTTGAATTCATA 3840  
QY 3841 AACTCATGTATGACATACATCAGAAATGCTGGGAAAGAAATTCATTTATGGAGGAG 3900  
DB 3841 AACTCATGTATGACATACATCAGAAATGCTGGGAAAGAAATTCATTTATGGAGGAG 3900  
QY 3901 AGAGCAACCCAGCAGTCACTACTCTAGAAATGCTTTGGGCTCTTAAACGGGGTCCAT 3960  
DB 3901 AGAGCAACCCAGCAGTCACTACTCTAGAAATGCTTTGGGCTCTTAAACGGGGTCCAT 3960  
QY 3961 CTGTTACTGAAAAGGAGGTGCTGAGTGCATCTTTTGGCCAAAGAAACAGAGAGGTGAAA 4020  
DB 3961 CTGTTACTGAAAAGGAGGTGCTGAGTGCATCTTTTGGCCAAAGAAACAGAGAGGTGAAA 4020  
QY 4021 TAGAAAAATTAATGCGCATGTTATTCGGCTGTGTCAGAAATCTACTGCTTTAACCCAGC 4080  
DB 4021 TAGAAAAATTAATGCGCATGTTATTCGGCTGTGTCAGAAATCTACTGCTTTAACCCAGC 4080  
QY 4081 ACAGGGAAAAACCCATAGAACTCTCAGGAGAGCCCTAGAGCCACTTTTCATGGATCCAG 4140  
DB 4081 ACAGGGAAAAACCCATAGAACTCTCAGGAGAGCCCTAGAGCCACTTTTCATGGATCCAG 4140  
QY 4141 ACTTGGCATATGGAACCTTATACAGAACTGTGCTCATGTAATGCAAGGAGTGTCTGGC 4200  
DB 4141 ACTTGGCATATGGAACCTTATACAGAACTGTGCTCATGTAATGCAAGGAGTGTCTGGC 4200  
QY 4201 AGAAGTATTTGAAAGCTGTACAGTGTGCTCAGCAGGCAATTCATGTTGACCTTTTGG 4260  
DB 4201 AGAAGTATTTGAAAGCTGTACAGTGTGCTCAGCAGGCAATTCATGTTGACCTTTTGG 4260  
QY 4261 ACTTGGAAAGTGGAGAAATCTTTGCGCTCTTTCGAAATCTCTGTCGCAATCTGTGATCC 4320  
DB 4261 ACTTGGAAAGTGGAGAAATCTTTGCGCTCTTTCGAAATCTCTGTCGCAATCTGTGATCC 4320  
QY 4321 CCATTTATTTCTTTGCAACTCAAAAGATAAAGAGTGAAGTGCAGATGCTTTTGTCTCAAC 4380  
DB 4321 CCATTTATTTCTTTGCAACTCAAAAGATAAAGAGTGAAGTGCAGATGCTTTTGTCTCAAC 4380  
QY 4381 TTTTGCACCTTGGCAGCGTGGGATACAGACTGTTTCTGGCCAGAAATATCAGGTTATATATA 4440

Db 4381 TTTTGA CCGTGG CACGGTGG GATA CAGACT GTTCTGG CCGCAG AATAT CAGGTATATATAA 4440  
Qy 4441 GACATGCTAAAGGAGAAAACCCAAATTCCTATTTTCTTTTAATCAAGGAATGGAGATCTA 4500  
Db 4441 GACATGCTAAAGGAGAAAACCCAAATTCCTATTTTCTTTAAATCAAGGAATGGAGATCTA 4500  
Qy 4501 CTTTGGAGTTCCATTCGATCCGATGTTTGGGTTGGAGTCTTCGATTAATTAATCAATA 4560  
Db 4501 CTTTGGAGTTCCATTCGATCCGATGTTTGGGTTGGAGTCTTCGATTAATTAATCAATA 4560  
Qy 4561 GCATCAAGGAAATGGTATTTCTCTTTGGCCACAACAATTTATAGAAATGGAAATGGC 4620  
Db 4561 GCATCAAGGAAATGGTATTTCTCTTTGGCCACAACAATTTATAGAAATGGAAATGGC 4620  
Qy 4621 CACCTGATGAAGAGGATCCTCGAGTCCCATGCTGAGCTGGAGCACCTGGCTTTCACTA 4680  
Db 4621 CACCTGATGAAGAGGATCCTCGAGTCCCATGCTGAGCTGGAGCACCTGGCTTTCACTA 4680  
Qy 4681 TCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGAACCTCTGTTTGGAGCACTTCAAA 4740  
Db 4681 TCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGAACCTCTGTTTGGAGCACTTCAAA 4740  
Qy 4741 ATAGCAGCAATAATGGTCTGAAAGCAATTAATGTCAGTTTGGCAGTTCACAGAGGATTAACCT 4800  
Db 4741 ATAGCAGCAATAATGGTCTGAAAGCAATTAATGTCAGTTTGGCAGTTCACAGAGGATTAACCT 4800  
Qy 4801 GTCCCTCAGGTCCTGTATACAGAAACATCTGGTTCCTATCATAGTGTCTTCTTCAACA 4860  
Db 4801 GTCCCTCAGGTCCTGTATACAGAAACATCTGGTTCCTATCATAGTGTCTTCTTCAACA 4860  
Qy 4861 TAAATACAGAGATACACCATGCTTCTGTATGGGATGACCTGTGTATAGATCTGTTTGGTGGTG 4920  
Db 4861 TAAATACAGAGATACACCATGCTTCTGTATGGGATGACCTGTGTATAGATCTGTTTGGTGGTG 4920  
Qy 4921 CTGTGTTAGCATTCCTCATCTGTTATGGGATGACCTGTGTATAGATCTGTTTGGTGGTG 4980  
Db 4921 CTGTGTTAGCATTCCTCATCTGTTATGGGATGACCTGTGTATAGATCTGTTTGGTGGTG 4980  
Qy 4981 TTAGTTCTTCTTATACCAACCTTTATCTCTTCATTTGATCACCATGAGCAACATGCTTC 5040  
Db 4981 TTAGTTCTTCTTATACCAACCTTTATCTCTTCATTTGATCACCATGAGCAACATGCTTC 5040  
Qy 5041 AGATACATTAAGTACAGACAGGCTTACCCCTGCTCAGGTTTCAAGAGACAGTGAAG 5100  
Db 5041 AGATACATTAAGTACAGACAGGCTTACCCCTGCTCAGGTTTCAAGAGACAGTGAAG 5100  
Qy 5101 AGGCTCATTCGSCATCTTCTTCTTTCAGAAATTTCTCAATATACAAAGTGGCTCAATG 5160  
Db 5101 AGGCTCATTCGSCATCTTCTTCTTTCAGAAATTTCTCAATATACAAAGTGGCTCAATG 5160  
Qy 5161 GGTGTGATATTCCTGGCTGGTATTTTGGGTCTCACTGAAGAAATGGCATCACCCCTTATC 5220  
Db 5161 GGTGTGATATTCCTGGCTGGTATTTTGGGTCTCACTGAAGAAATGGCATCACCCCTTATC 5220  
Qy 5221 TTCGCTGTGCTGATTTTTCATTTTCACTATTTTCACTATTTTCACTATTTTCACTATTTT 5280  
Db 5221 TTCGCTGTGCTGATTTTTCATTTTCACTATTTTCACTATTTTCACTATTTTCACTATTTT 5280  
Qy 5281 ATACCAAATTCGAGAGAGAGTACAGTGCACTCTGTAGCTATCTATCTTTTACCTACAA 5340  
Db 5281 ATACCAAATTCGAGAGAGAGTACAGTGCACTCTGTAGCTATCTATCTTTTACCTACAA 5340  
Qy 5341 ATTTGTTCTGCTCTTCCAGGAATATTTGGGATCTGTAAGGCCCTTGTCTCCAGAGTGGT 5400  
Db 5341 ATTTGTTCTGCTCTTCCAGGAATATTTGGGATCTGTAAGGCCCTTGTCTCCAGAGTGGT 5400  
Qy 5401 GTGCAGATCTCGCTTACTAACTGTTTGAAGCAAAAAACCCGTGTGCTCAGTACCTTA 5460  
Db 5401 GTGCAGATCTCGCTTACTAACTGTTTGAAGCAAAAAACCCGTGTGCTCAGTACCTTA 5460  
Qy 5461 GAAAAAGAAATAGTTTGTATAGAGCTTCTGTATGATCTATATAGTGTGCTCTCTGCAATTTGATTT 5520

Db 5461 GAAAAAGAAATAGTTTGTATAGAGCTTCTGTATGACTATAGCTGCTCTCTGAATCAAGCTT 5520  
Qy 5521 CTCAATTCAGGTGCCCCAGGCTCTGCAGATGATGAGCGAAGCATCCCTGCTCTCTGCTTT 5580  
Db 5521 CTCAATTCAGGTGCCCCAGGCTCTGCAGATGATGAGCGAAGCATCCCTGCTCTCTGCTTT 5580  
Qy 5581 TCTGTGGGGCTATACATATATTTCTCAGAAACATTTTCTGCCAGGAAATTTGTGAACGGGAAG 5640  
Db 5581 TCTGTGGGGCTATACATATATTTCTCAGAAACATTTTCTGCCAGGAAATTTGTGAACGGGAAG 5640  
Qy 5641 AGGTGTGAGCTTGCAATTTTCAAGCACTTCACTGTGAGCGCGAGTCTGCAATTTTCTTAA 5700  
Db 5641 AGGTGTGAGCTTGCAATTTTCAAGCACTTCACTGTGAGCGCGAGTCTGCAATTTTCTTAA 5700  
Qy 5701 AAATCAGAGATCCGAGTGTCTGCTGCTGAAGTAAAGCCAGAGGCTGTGCCCTATCCAG 5760  
Db 5701 AAATCAGAGATCCGAGTGTCTGCTGCTGAAGTAAAGCCAGAGGCTGTGCCCTATCCAG 5760  
Qy 5761 CTCTTACTTGGATGAATATGGGAAACAGACCCCTGCCCTGAAAGGGGCAACCCCTTTC 5820  
Db 5761 CTCTTACTTGGATGAATATGGGAAACAGACCCCTGCCCTGAAAGGGGCAACCCCTTTC 5820  
Qy 5821 ATTTATCTGTGAGCGGTATCGGAAGCTTCCATTTTGGTCTGGCAACAACA CTGCAATTTATAG 5880  
Db 5821 ATTTATCTGTGAGCGGTATCGGAAGCTTCCATTTTGGTCTGGCAACAACA CTGCAATTTATAG 5880  
Qy 5881 AAGAGATGCTTGGAGCCAGAGCACTAATCAGATGTTTATTTGGATTTCAACTGGCAGTTAC 5940  
Db 5881 AAGAGATGCTTGGAGCCAGAGCACTAATCAGATGTTTATTTGGATTTCAACTGGCAGTTAC 5940  
Qy 5941 TGTGAGCTTCCAACTCTGCTCCTCAAGACAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000  
Db 5941 TGTGAGCTTCCAACTCTGCTCCTCAAGACAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000  
Qy 6001 AAAATTAAGGAAACTTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTTTTGTCTCCATGTC 6060  
Db 6001 AAAATTAAGGAAACTTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTTTTGTCTCCATGTC 6060  
Qy 6061 CAGGTTCACTTACATCAATTAATTTCTTAAATGAGTATTTGCTTTCAATTAGCAACA 6120  
Db 6061 CAGGTTCACTTACATCAATTAATTTCTTAAATGAGTATTTGCTTTCAATTAGCAACA 6120  
Qy 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180  
Db 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180  
Qy 6181 ATTTATCCCTTCTCATATTTGCTCATTTTCAATTTTATTTTCACTGCTGAGTGAAG 6240  
Db 6181 ATTTATCCCTTCTCATATTTGCTCATTTTCAATTTTATTTTCACTGCTGAGTGAAG 6240  
Qy 6241 TCAGTCAACAGTTGTAGACATTTATGTTGTTGTTAACTCTTCTGCAATTTTGTATTT 6300  
Db 6241 TCAGTCAACAGTTGTAGACATTTATGTTGTTGTTAACTCTTCTGCAATTTTGTATTT 6300  
Qy 6301 GGTGTTTTT 6308  
Db 6301 GGTGTTTTT 6308

RESULT 2  
US-10-758-636A-1  
; Sequence 1, Application US/10758636A  
; Publication No. US2005009876A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966C  
; CURRENT APPLICATION NUMBER: US/10/758,636A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; FILING DATE: 2000-03-08

```
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 6308
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (696)..(5942)
US-10-758-636A-1

Query Match      100.0%; Score 6308; DB 21; Length 6308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAAGAATTCCGCGACGAGGGGAAAGCTGAGCCAGGAAACCAAAATTAATCTGCTTTTACCT 60
DB 1 GCCAAGAATTCCGCGACGAGGGGAAAGCTGAGCCAGGAAACCAAAATTAATCTGCTTTACCT 60

QY 61 CATTTGTGAAGCAAGCGTCAAAACAGCTTCAACTATCTTTGAACAAGAGAACTTACCT 120
DB 61 CATTTGTGAAGCAAGCGTCAAAACAGCTTCAACTATCTTTGAACAAGAGAACTTACCT 120

QY 121 CCAAGGCTTATCATCTGCTTCCACTTATCCAAAGCTGATGCGCCACTGCGCTGTC 180
DB 121 CCAAGGCTTATCATCTGCTTCCACTTATCCAAAGCTGATGCGCCACTGCGCTGTC 180

QY 181 CGCACTTGAACACCGCCAGCCCACTACTGCTCCCACTACCACTGTTCTCCCACTCT 240
DB 181 CGCACTTGAACACCGCCAGCCCACTACTGCTCCCACTACCACTGTTCTCCCACTCT 240

QY 241 GATAGCTGCTTGTGCTGCCATCTATTCGGCTTCTGCTGCTTCTGAGTAAATGATACA 300
DB 241 GATAGCTGCTTGTGCTGCCATCTATTCGGCTTCTGCTGCTTCTGAGTAAATGATACA 300

QY 301 CACCTGTGAACCAACCATTTCTACTTCTGCTGCTATGATTTGACTCTAGCTGGATC 360
DB 301 CACCTGTGAACCAACCATTTCTACTTCTGCTGCTATGATTTGACTCTAGCTGGATC 360

QY 361 CCGAGCTTTTGTACATGTGCAAGTGCAGGCGCCAGGGGTAGAAATCTTAAATAAGAGAT 420
DB 361 CCGAGCTTTTGTACATGTGCAAGTGCAGGCGCCAGGGGTAGAAATCTTAAATAAGAGAT 420

QY 421 GTATGCAACAGTTCCAGACACAAACCCAGATATACCACTTACGCTACCAAGAGCTAC 480
DB 421 GTATGCAACAGTTCCAGACACAAACCCAGATATACCACTTACGCTACCAAGAGCTAC 480

QY 481 GCCTGATAAATTAGAGGGGAAAAAATAATCTCCAGTCCCTTCACTGCTGAGCGTTGCT 540
DB 481 GCCTGATAAATTAGAGGGGAAAAAATAATCTCCAGTCCCTTCACTGCTGAGCGTTGCT 540

QY 541 TCCGGGAAGCGGCGGGAAGCACTCCTCGAGTCTGCGTCAAAACCGGACTTTCAGGGGCG 600
DB 541 TCCGGGAAGCGGCGGGAAGCACTCCTCGAGTCTGCGTCAAAACCGGACTTTCAGGGGCG 600

QY 601 TCGTAAAAAGTGTGCTCCTGCTCTCCGACGGGCAAGAGTTTCGCTTTCGCTGCGCG 660
DB 601 TCGTAAAAAGTGTGCTCCTGCTCTCCGACGGGCAAGAGTTTCGCTTTCGCTGCGCG 660

QY 661 GGGTTCGCAACTGACAGCGTCAGTTTCCTCAAGATGCGGACGAGAGGCTGAGGTA 720
DB 661 GGGTTCGCAACTGACAGCGTCAGTTTCCTCAAGATGCGGACGAGAGGCTGAGGTA 720

QY 721 CTGAGAGATGGAATCAGCGGGAGTTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780
DB 721 CTGAGAGATGGAATCAGCGGGAGTTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780

QY 781 GGGATCAGCAGTTGATTTTATATCTGCTTTCTGCATCAITTTGCACAATTTGGTCCAG 840
DB 781 GGGATCAGCAGTTGATTTTATATCTGCTTTCTGCATCAITTTGCACAATTTGGTCCAG 840

QY 841 AAATTTACTTGTGCTGAATGGAACCCAGACTTGGAAAAAGCAGAGGAAAGTGTACAAATGT 900
DB 841 AAATTTACTTGTGCTGAATGGAACCCAGACTTGGAAAAAGCAGAGGAAAGTGTACAAATGT 900

841 AAATTTACTTGTGCTGAATGGAACCCAGACTTGGAAAAAGCAGAGGAAAGTGTACAAATGT 900
901 CAATATTTCACTCCACTCGGATGTTACTTATTTTGGAGAGATCCAGATATTTCTCTTAGAGA 960
901 CAATATTTCACTCCACTCGGATGTTACTTATTTTGGAGAGATCCAGATATTTCTCTTAGAGA 960
961 AATTGAAGCAGATGAGGAGCAATTTAGCTTTTGTGGAGGGTTTTTCAAAAGTGGAGAGACAA 1020
961 AATTGAAGCAGATGAGGAGCAATTTAGCTTTTGTGGAGGGTTTTTCAAAAGTGGAGAGACAA 1020
1021 CCTATTTCTGAGGAGATTTGCAATTTGATCCAAATGTTACTCTGTATGGAAGTCTTCTCC 1080
1021 CCTATTTCTGAGGAGATTTGCAATTTGATCCAAATGTTACTCTGTATGGAAGTCTTCTCC 1080
1081 AGGACAGTGTTCATAAAAATCATCGTTTACAGATGCATATCTTCTACTGAGGAGGGTTCT 1140
1081 AGGACAGTGTTCATAAAAATCATCGTTTACAGATGCATATCTTCTACTGAGGAGGGTTCT 1140
1141 GTGACTGTGGAGACACAGAGGATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200
1141 GTGACTGTGGAGACACAGAGGATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200
1201 GAAGAGCAGGTACTATTAAGAGAGATTCACGCTGTCGTTGAATCAAGAGGTAATTTGTC 1260
1201 GAAGAGCAGGTACTATTAAGAGAGATTCACGCTGTCGTTGAATCAAGAGGTAATTTGTC 1260
1261 AAGCAGGAGAAATATTTCTTTCAGTGATAAATATGTCGTAGAAATGACTATATGGGAG 1320
1261 AAGCAGGAGAAATATTTCTTTCAGTGATAAATATGTCGTAGAAATGACTATATGGGAG 1320
1321 AGGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGAGAAATCAAAAGATATATTTGTG 1380
1321 AGGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGAGAAATCAAAAGATATATTTGTG 1380
1381 TCCTTTTCAATGATGAAACCACTATCATATGACCACTATATACAGCTTACAAAGAGCTC 1440
1381 TCCTTTTCAATGATGAAACCACTATCATATGACCACTATATACAGCTTACAAAGAGCTC 1440
1441 TTGACTGTGAGCTCGCAGAGGCCCACTGTCATACCACTGCCATTCACAAAGAGGGTCTGTC 1500
1441 TTGACTGTGAGCTCGCAGAGGCCCACTGTCATACCACTGCCATTCACAAAGAGGGTCTGTC 1500
1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTGCAGGAAAGCAAGAGATATATAAGAGTTC 1560
1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTGCAGGAAAGCAAGAGATATATAAGAGTTC 1560
1561 ATTCAAGAAATGCTCTCAACATCCATTCATGATGAGATTAACCTACAGAGATTTAGG 1620
1561 ATTCAAGAAATGCTCTCAACATCCATTCATGATGAGATTAACCTACAGAGATTTAGG 1620
1621 CTCATCAGAAATTTGCTTTGCTGCTTGGTCTTGGATGAAACAAATATATGAGCTATTCAA 1680
1621 CTCATCAGAAATTTGCTTTGCTGCTTGGTCTTGGATGAAACAAATATATGAGCTATTCAA 1680
1681 GTGACTTTAGCAGATCTTTTGCAGCATGCTTTAGAGAGAAACCTGACTCGGAGATTC 1740
1681 GTGACTTTAGCAGATCTTTTGCAGCATGCTTTAGAGAGAAACCTGACTCGGAGATTC 1740
1741 CCTGCTCATAGCAGGTTAAATGCTTTGGGATGCAAGCTTTTATAAGGTTGCCGTAAAGA 1800
1741 CCTGCTCATAGCAGGTTAAATGCTTTGGGATGCAAGCTTTTATAAGGTTGCCGTAAAGA 1800
1801 TCCTTCTAGAAATTTGATCTTCAAGAGTTTATTTTATGAGATGGAATACAAAAACCTCTTTG 1860
1801 TCCTTCTAGAAATTTGATCTTCAAGAGTTTATTTTATGAGATGGAATACAAAAACCTCTTTG 1860
1861 CTATGGAATTTGTGAAGTATTTAAACAACTGCGAGAAAGATATATATCATGATGATCATG 1920
1861 CTATGGAATTTGTGAAGTATTTAAACAACTGCGAGAAAGATATATATCATGATGATCATG 1920
1921 ACAGAGTATCTCTATTAACCTGCACTTTTCAGTTTCAGATGTTTACTGTTCTTCTTCTGCTC 1980
1921 ACAGAGTATCTCTATTAACCTGCACTTTTCAGTTTCAGATGTTTACTGTTCTTCTTCTGCTC 1980
```







RESULT 3  
US-10-287-218-22  
; Sequence 22, Application US/10287218  
; Publication NO. US20030198975A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.  
; APPLICANT: BATRA, Sajeesh; BAUGHN, Mariah R.  
; APPLICANT: BECHA, Shanya D.; BOROWSKI, Mark L.  
; APPLICANT: BUFORD, Neil; DING, Li  
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.  
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.  
; APPLICANT: GRIFFIN, Jennifer A.; HAFALIA, April J.A.  
; APPLICANT: HONCHELLI, Cynthia D.; LAL, Preti G.  
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.  
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi  
; APPLICANT: REDDY, Roopa; SANJANWALA, Madhu, M.  
; APPLICANT: TANG, Y. Tom; WALIA, Narinder K.  
; APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.  
; APPLICANT: XU, Yuming; YANG, Junming  
; APPLICANT: YAO, Monique G.; YUE, Henry  
; APPLICANT: ZEBARJADIAN, Yeganeh  
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH  
; FILE REFERENCE: PI-0417 USA  
; CURRENT APPLICATION NUMBER: US/10/287,218  
; CURRENT FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: PCT/US02/11152  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: US 60/349,705  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: US 60/295,263  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/295,340  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/293,727  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 60/291,846  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/291,662  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: US 60/287,228  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/286,820  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/283,294  
; PRIOR FILING DATE: 2001-04-11  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 22  
; LENGTH: 7742  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20030198975A1 1351608CB1  
US-10-287-218-22

Query Match 88.6%; Score 5592; DB 16; Length 7742;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 717 GGTACTGAGAGATGGAATCAGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCT 776  
Db 1 GGTACTGAGAGATGGAATCAGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCT 60

QY 777 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTGATCATTTGGCACAATTGGTG 836  
Db 61 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTGATCATTTGGCACAATTGGTG 120

QY 837 CCAGAAATTTACTTTGCTGAAATGGACCAGACTTGGAAAGCAGGAGAAAGTGTACAA 896

Db 121 CCAGAAATTTACTTTGCTGAAATGGACCAGACTTGGAAAGCAGGAGAAAGTGTACAA 180  
QY 897 ATGTCAATATTTCACTCCACTGGATGTACTTATTTGGAGAGATCCAGATATTTGCTTA 956  
Db 181 ATGTCAATATTTCACTCCACTGGATGTACTTATTTGGAGAGATCCAGATATTTGCTTA 240  
QY 957 GAGAAATTTGAAGCACAGTGGAGCAATTTTCAGCTTTGTGGAGGGTTTTCAAAAGTGGAGAG 1016  
Db 241 GAGAAATTTGAAGCACAGTGGAGCAATTTTCAGCTTTGTGGAGGGTTTTCAAAAGTGGAGAG 300  
QY 1017 ACAACCTATTTCTCGAGGATTTGCAATTGATCCAAATGTGTAATCTCTGTATGAGACTGC 1076  
Db 301 ACAACCTATTTCTCGAGGATTTGCAATTGATCCAAATGTGTAATCTCTGTATGAGACTGC 360  
QY 1077 TTCCAGGACAGTGTTCATTAATAATCATCGTTTACAGATGCATCTTCTCTGAGGAGGG 1136  
Db 361 TTCCAGGACAGTGTTCATTAATAATCATCGTTTACAGATGCATCTTCTCTGAGGAGGG 420  
QY 1137 TTCTGTGACTGTGGAGACACAGAGGCATGGAAACTGGCCCTTTTGTGTAAATCATGAA 1196  
Db 421 TTCTGTGACTGTGGAGACACAGAGGCATGGAAACTGGCCCTTTTGTGTAAATCATGAA 480  
QY 1197 CCTGGAAGAGCAGTACTATATAAAGAGAAATTCACGCTGTCCGTTGAATGAAGAGTAAT 1256  
Db 481 CCTGGAAGAGCAGTACTATATAAAGAGAAATTCACGCTGTCCGTTGAATGAAGAGTAAT 540  
QY 1257 GTCCAGGCCAGGAAATATTTCTTCAGTGTATAAATATGTCTGTAGAAATCAGTATATGG 1316  
Db 541 GTCCAGGCCAGGAAATATTTCTTCAGTGTATAAATATGTCTGTAGAAATCAGTATATGG 600  
QY 1317 GAAGAGGAAAAAGAACTGCTCCCTGAACTCCAGATAAGGAGAAAAATGAAGATACTAT 1376  
Db 601 GAAGAGGAAAAAGAACTGCTCCCTGAACTCCAGATAAGGAGAAAAATGAAGATACTAT 660  
QY 1377 TGTGTCTTTTCAATGATGAACACCAATTCATATGACACCGTCTATATACAGCTTACAAAGA 1436  
Db 661 TGTGTCTTTTCAATGATGAACACCAATTCATATGACACCGTCTATATACAGCTTACAAAGA 720  
QY 1437 GCTCTTGACTGTGAGCTGGAGAGGCCAGTGTGATACCACCTGCTCCATTTGACAAAGAGGT 1496  
Db 721 GCTCTTGACTGTGAGCTGGAGAGGCCAGTGTGATACCACCTGCTCCATTTGACAAAGAGGT 780  
QY 1497 CGTGGGCTGTAAAAGCGGAGCTTATGCTCTTGCAGGAAGCAAAAGAAATATAAG 1556  
Db 781 CGTGGGCTGTAAAAGCGGAGCTTATGCTCTTGCAGGAAGCAAAAGAAATATAAG 840  
QY 1557 AGTCATTCAGAAAAATGCTCTCAACATCCACTTCAATGTAGAAGTATTAACCTCAGAGATT 1616  
Db 841 AGTCATTCAGAAAAATGCTCTCAACATCCACTTCAATGTAGAAGTATTAACCTCAGAGATT 900  
QY 1617 ATGGCTCATCAGAAATTTGCTTTGGCTTTGGTCTTGGATGAACAAATTTATGAGCTAT 1676  
Db 901 ATGGCTCATCAGAAATTTGCTTTGGCTTTGGTCTTGGATGAACAAATTTATGAGCTAT 960  
QY 1677 TCAAGTGAATTTAGCGAGATCTTTTGGCAACATGCTTTAGAGAAAGAACCTGACTCGGAG 1736  
Db 961 TCAAGTGAATTTAGCGAGATCTTTTGGCAACATGCTTTAGAGAAAGAACCTGACTCGGAG 1020  
QY 1737 AATCCCTGTCTCATAGCAGGTTAATGCTTTGGATGCAAGCTTTTATAAGGTGCCCCGT 1796  
Db 1021 AATCCCTGTCTCATAGCAGGTTAATGCTTTGGATGCAAGCTTTTATAAGGTGCCCCGT 1080  
QY 1797 AAGATCTTTCATGAATTTGATCTTCAGGAGTGTGATTTTATGGAGATGGATACAAAACTC 1856  
Db 1081 AAGATCTTTCATGAATTTGATCTTCAGGAGTGTGATTTTATGGAGATGGATACAAAACTC 1140  
QY 1857 TTTGCTATGGAATTTTGTGAAGTATTTATAAACAACCTGCAAGAAAGATATATCAATGATCAT 1916  
Db 1141 TTTGCTATGGAATTTTGTGAAGTATTTATAAACAACCTGCAAGAAAGATATATCAATGATCAT 1200  
QY 1917 CATGACAGAAAGTATCTCTATAAATGCACTTTTCAGTTCAGATGTTTACCTGTTCTTACTCTG 1976

Db 1201 CATGACAGAGATCTCTATTAACGACCTTTTCAGTTTCAGATGTTTACTGTTCTCTACTCTG 1260  
Qy 1977 GCTCGACATCTTATTGAAGACAGAAATGTTATCTCTGTCATTTACTGAAATCTCTGCTAGAA 2036  
Db 1261 GCTCGACATCTTATTGAAGACAGAAATGTTATCTCTGTCATTTACTGAAATCTCTGCTAGAA 1320  
Qy 2037 GTTCTTACCTGAGTACTTGGACAGGAAACAAATAATTCACCTTCAGGGTTTATAGCCAGGAC 2096  
Db 1321 GTTCTTACCTGAGTACTTGGACAGGAAACAAATAATTCACCTTCAGGGTTTATAGCCAGGAC 1380  
Qy 2097 AAATTTGGGAAGAGTATATGTCAGTAAATATGTCAGCTTAAAGTATATCTGATCAGCAAAACC 2156  
Db 1381 AAATTTGGGAAGAGTATATGTCAGTAAATATGTCAGCTTAAAGTATATCTGATCAGCAAAACC 1440  
Qy 2157 ACAATATGACAGAAAGATTAAAGATGCAATGCTTGAAGGTTTTCGATCTTTTGAAG 2216  
Db 1441 ACAATATGACAGAAAGATTAAAGATGCAATGCTTGAAGGTTTTCGATCTTTTGAAG 1500  
Qy 2217 ATTCTTACCTGTATGTCAGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTCGA 2276  
Db 1501 ATTCTTACCTGTATGTCAGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTCGA 1560  
Qy 2277 GTGGATCTGATTTGGGAGGCTGCCATTTGCTATACAGATGCAATTCGAAGAAATATTTTACTC 2336  
Db 1561 GTGGATCTGATTTGGGAGGCTGCCATTTGCTATACAGATGCAATTCGAAGAAATATTTTACTC 1620  
Qy 2337 ATGTTCCAGAGTGCTGTTGATGAGAACTCTTACTTGTGGCTTATAAGAAATGT 2396  
Db 1621 ATGTTCCAGAGTGCTGTTGATGAGAACTCTTACTTGTGGCTTATAAGAAATGT 1680  
Qy 2397 CACAAGCTGTGATGAGTGACGATACAGTTTCTATCTAGTAGCAAGCAGTAGTACAA 2456  
Db 1691 CACAAGCTGTGATGAGTGACGATACAGTTTCTATCTAGTAGCAAGCAGTAGTACAA 1740  
Qy 2457 TCGTGTGACATAGTTTGGAAACAAAGCTCTACAGATGATCTGAGGATCTTGTAGACATA 2516  
Db 1741 TCGTGTGACATAGTTTGGAAACAAAGCTCTACAGATGATCTGAGGATCTTGTAGACATA 1800  
Qy 2517 CATCTGCCACTCTTAGACCTCTGCTGCTTCATGTAAGTTAAGCAGGCTGGGTGCT 2576  
Db 1801 CATCTGCCACTCTTAGACCTCTGCTGCTTCATGTAAGTTAAGCAGGCTGGGTGCT 1860  
Qy 2577 GTTTCAGACTGTCATGAATTTGTGTTGCTTGTGAGGACTTTCAGTAGAGGTACTAGTGAA 2636  
Db 1861 GTTTCAGACTGTCATGAATTTGTGTTGCTTGTGAGGACTTTCAGTAGAGGTACTAGTGAA 1920  
Qy 2637 TATCTCTTTACGTTGCTGCTGTTGCTGCCAGGTTGTTGCTGAGATGTTGGCGAAGAAAT 2696  
Db 1921 TATCTCTTTACGTTGCTGCTGTTGCTGCCAGGTTGTTGCTGAGATGTTGGCGAAGAAAT 1980  
Qy 2697 GGACTGCTCTTATTAGCCAGGTGTTTATTACCAAGATGTTAAGTGACAGAGAGAAATG 2756  
Db 1981 GGACTGCTCTTATTAGCCAGGTGTTTATTACCAAGATGTTAAGTGACAGAGAGAAATG 2040  
Qy 2757 TATGATAAGATATCATCATGCTTCAGATTGGTGCATCTTTAATGGATGCCAATAAGTTC 2816  
Db 2041 TATGATAAGATATCATCATGCTTCAGATTGGTGCATCTTTAATGGATGCCAATAAGTTC 2100  
Qy 2817 TTGTTTACCTGTTACAGAGGTATGAACTTGCAGGCTTTTAAACAAGACCATATCTACA 2876  
Db 2101 TTGTTTACCTGTTACAGAGGTATGAACTTGCAGGCTTTTAAACAAGACCATATCTACA 2160  
Qy 2877 AAAGACACAGGATTTGATTTAAACAATATATACATAAGAGAAATGCTTCAGGTCTCTC 2936  
Db 2161 AAAGACACAGGATTTGATTTAAACAATATATACATAAGAGAAATGCTTCAGGTCTCTC 2220  
Qy 2937 ATCTATATGTTGGGTGAGGTTATGTAATCTGAGTGGGAATGTAACAAAGAGAGGTTC 2996  
Db 2221 ATCTATATGTTGGGTGAGGTTATGTAATCTGAGTGGGAATGTAACAAAGAGAGGTTC 2280  
Qy 2997 ACNATGAGAGAAATCAATTCATCTGCTTGTGATTTGAACCCCATGCCACAGTGCCATTTGCC 3056  
Db 2281 ACNATGAGAGAAATCAATTCATCTGCTTGTGATTTGAACCCCATGCCACAGTGCCATTTGCC 2340

Qy 3057 AAAAATTTTACCTGAGAAATGAATAATGAATACTGCTTTAGAGAAATGTCATATAAACAAGTG 3116  
Db 2341 AAAAATTTTACCTGAGAAATGAATAATGAATACTGCTTTAGAGAAATGTCATATAAACAAGTG 2400  
Qy 3117 GCCACATTTTAAAGAAACCAAGGTGTATCAGGCCATGAGTTTATGAACATAAAGATGAATCA 3176  
Db 2401 GCCACATTTTAAAGAAACCAAGGTGTATCAGGCCATGAGTTTATGAACATAAAGATGAATCA 2460  
Qy 3177 CTGAAGACTTTCAATATGTAATCTTTTATCAITTACTCTCAAAACCAGCATAGCAAGGCTGAA 3236  
Db 2461 CTGAAGACTTTCAATATGTAATCTTTTATCAITTACTCTCAAAACCAGCATAGCAAGGCTGAA 2520  
Qy 3237 CATATGACAGAGAAAGAGGAGAAACCAAGAAACCAAGATGAAGCATTTGCCGCCACACCA 3296  
Db 2521 CATATGACAGAGAAAGAGGAGAAACCAAGAAACCAAGATGAAGCATTTGCCGCCACACCA 2580  
Qy 3297 CCTCTGAATTTCTGCCCTGCTTTTCAGCAAGAGTATTAACTTCTCAACTTGTGATATCATG 3356  
Db 2591 CCTCTGAATTTCTGCCCTGCTTTTCAGCAAGAGTATTAACTTCTCAACTTGTGATATCATG 2640  
Qy 3357 ATGTACATTTCTCAGGACCGTATTTTGAAGCGGCAATAGACACAGATTTCTAACTTTGTGACC 3416  
Db 2641 ATGTACATTTCTCAGGACCGTATTTTGAAGCGGCAATAGACACAGATTTCTAACTTTGTGACC 2700  
Qy 3417 GAAGGATGCTCCAAATGCGCTTTTCATATTTCTGGCATTTGGGTTTACTAGAGAGAGCA 3476  
Db 2701 GAAGGATGCTCCAAATGCGCTTTTCATATTTCTGGCATTTGGGTTTACTAGAGAGAGCA 2760  
Qy 3477 CAGCTTCAAAAGCTCTTGAAGAGAGTAACTTTGATTTTATCATTAAGCTTTCAAGA 3536  
Db 2761 CAGCTTCAAAAGCTCTTGAAGAGAGTAACTTTGATTTTATCATTAAGCTTTCAAGA 2820  
Qy 3537 TTGGAGGTTCTCAGGACATGATATATAAATGCTTTTGGAAAACTCAAAAGGAATTTCCCGAG 3596  
Db 2821 TTGGAGGTTCTCAGGACATGATATATAAATGCTTTTGGAAAACTCAAAAGGAATTTCCCGAG 2880  
Qy 3597 TTAGAGGCCAGAGAGGACATGATAACGTTGATATTTTCAAGATGTTTGAACAGTGAAGCA 3656  
Db 2881 TTAGAGGCCAGAGAGGACATGATAACGTTGATATTTTCAAGATGTTTGAACAGTGAAGCA 2940  
Qy 3657 TTAAGAGAAATCTTGTGTTTAAATTTGTAGCAACCAATCAGGATCGGAATCTTATTAAGAT 3716  
Db 2941 TTAAGAGAAATCTTGTGTTTAAATTTGTAGCAACCAATCAGGATCGGAATCTTATTAAGAT 3000  
Qy 3717 GATGAGATTACTCATGATAGAAAAAGACAGAAACGAAAAAGAAAGCTGAAAGCTGCTAGG 3776  
Db 3001 GATGAGATTACTCATGATAGAAAAAGACAGAAACGAAAAAGAAAGCTGAAAGCTGCTAGG 3060  
Qy 3777 CTAATCGCCAGAGAGATCATGGCTCAGATGCTGCTTTACAGAAAAAATCTTCAITGAAAT 3836  
Db 3061 CTAATCGCCAGAGAGATCATGGCTCAGATGCTGCTTTACAGAAAAAATCTTCAITGAAAT 3120  
Qy 3837 CATAACTCATGATGATGACAAATACATCAAGAAATGCTGGGAAGAGATTCATTTATGAG 3896  
Db 3121 CATAACTCATGATGATGACAAATACATCAAGAAATGCTGGGAAGAGATTCATTTATGAG 3180  
Qy 3897 GAAGAGAGACCCACAGCAGTCAAGTCACTCTAGAAATTTGCTTTGGGCTCTAAACGGGT 3956  
Db 3181 GAAGAGAGACCCACAGCAGTCAAGTCACTCTAGAAATTTGCTTTGGGCTCTAAACGGGT 3240  
Qy 3957 CCATCTGTTACTGAAAAAGGAGGTCTGACGTGCTCTGCTTTCAGAAAAAATCTTCAITGAAAT 4016  
Db 3241 CCATCTGTTACTGAAAAAGGAGGTCTGACGTGCTCTTTCAGAAAAAATCTTCAITGAAAT 3300  
Qy 4017 AAAATAGAAAAATATGCTATGCTTATTCGGCTGCTGTCAGAAATCTACTGCTTAAACC 4076  
Db 3301 AAAATAGAAAAATATGCTATGCTTATTCGGCTGCTGTCAGAAATCTACTGCTTAAACC 3360  
Qy 4077 CAGCAGAGGGGAAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTTCATGAT 4136  
Db 3361 CAGCAGAGGGGAAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTTCATGAT 3420

QY 4137 CCAGACTGGCATAATGGAACCTTATACAGGAAGCTGTGGTCAATGTAAATGCAAGCAGTGTGC 4196  
DB 3421 CCAGACTTGGCATAATGGAACCTTATACAGGAAGCTGTGGTCAATGTAAATGCAAGCAGTGTGC 3480  
QY 4197 TGGCAGAGATATTTTGAAGCTGTACAGCTGAGCTCTCAGAGCGCATCAATGTTGACCTT 4256  
DB 3481 TGGCAGAGATATTTTGAAGCTGTACAGCTGAGCTCTCAGAGCGCATCAATGTTGACCTT 3540  
QY 4257 TTTGACTTGGAAAGTGGGAATATCTTTGGCCCTCTTTGCAATCTCTGTGCAATACTGTG 4316  
DB 3541 TTTGACTTGGAAAGTGGGAATATCTTTGGCCCTCTTTGCAATCTCTGTGCAATACTGTG 3600  
QY 4317 ATCCCACTTATTCCTTTGCAACTCAAAAGATATAACAGTGAAGTGAAGTGTCTGTGCT 4376  
DB 3601 ATCCCACTTATTCCTTTGCAACTCAAAAGATATAACAGTGAAGTGAAGTGTCTGTGCT 3660  
QY 4377 CAATCTTTGACCTGGCAGCGTGGATACAGACTGTCTGTGGCAGAAATATCAGGTTATAAT 4436  
DB 3661 CAATCTTTGACCTGGCAGCGTGGATACAGACTGTCTGTGGCAGAAATATCAGGTTATAAT 3720  
QY 4437 ATAAGACATGCTAAAGGGAAGAAACCAATTCCTATTTCTTTAATCAAGGAATGGAGAT 4496  
DB 3721 ATAAGACATGCTAAAGGGAAGAAACCAATTCCTATTTCTTTAATCAAGGAATGGAGAT 3780  
QY 4497 TCTACTTTGGAGTTCATTTCCATCTGAGTTTGTGGTGTGAGTCTTCGATTAATAATTTCA 4556  
DB 3781 TCTACTTTGGAGTTCATTTCCATCTGAGTTTGTGGTGTGAGTCTTCGATTAATAATTTCA 3840  
QY 4557 AATAGCATCAAGGAAATGGTTATCTCTTTTGCCACAACAATTTATAGAAATGGATGAAA 4616  
DB 3841 AATAGCATCAAGGAAATGGTTATCTCTTTTGCCACAACAATTTATAGAAATGGATGAAA 3900  
QY 4617 GTGCCACCTGATGAAGGGATCTCGAGTCCCATGCTGACCTGAGCAGCACCTGGCTTTC 4676  
DB 3901 GTGCCACCTGATGAAGGGATCTCGAGTCCCATGCTGACCTGAGCAGCACCTGGCTTTC 3960  
QY 4677 ACTATCCAGGCAATGAAATCTATTTGGAGATGAAGGAAACCTCTGTTTGGAGCACTT 4736  
DB 3961 ACTATCCAGGCAATGAAATCTATTTGGAGATGAAGGAAACCTCTGTTTGGAGCACTT 4020  
QY 4737 CAAATATAGCAGCATAAATGGTCTGAAAGCATTAATGCAATTTGCAAGTTGCGACAGAGGAT 4796  
DB 4021 CAAATATAGCAGCATAAATGGTCTGAAAGCATTAATGCAATTTGCAAGTTGCGACAGAGGAT 4080  
QY 4797 ACTGTCTCAGCTCTGATACAGAAATCTGTTGGTCTGCTCTATCAGTTGCTTCTCT 4856  
DB 4081 ACTGTCTCAGCTCTGATACAGAAACATCTGTTGGTCTGCTCTATCAGTTGCTTCTCT 4140  
QY 4857 AACATAAATCAGAAGATACACATGCTCTGCTATAGATCTGTTCATGTTTGGTG 4916  
DB 4141 AACATAAATCAGAAGATACACATGCTCTGCTATAGATCTGTTCATGTTTGGTG 4200  
QY 4917 GGTGCTGTGTAGCATCCCATCTGTATTTGGATGACCCCTGTTGATCTGAGCCTTCT 4976  
DB 4201 GGTGCTGTGTAGCATCCCATCTGTATTTGGATGACCCCTGTTGATCTGAGCCTTCT 4260  
QY 4977 TCAGTTAGTCTTCTATTAACACCTTATCTCTTCCATTTGATCAACCATGSCACATG 5036  
DB 4261 TCAGTTAGTCTTCTATTAACACCTTATCTCTTCCATTTGATCAACCATGSCACATG 4320  
QY 5037 CTTTCAGATACTACTTACAGTAGACACAGGCTTACCCCTTGCTCAGGTTCAAGAAGCAGT 5096  
DB 4321 CTTTCAGATACTACTTACAGTAGACACAGGCTTACCCCTTGCTCAGGTTCAAGAAGCAGT 4380  
QY 5097 GAAGAGGCTCAATTCGCACTCTTTCTTTTGAGAAATTTCTCAATATACAGTGGCTCC 5156  
DB 4381 GAAGAGGCTCAATTCGCACTCTTTCTTTTGAGAAATTTCTCAATATACAGTGGCTCC 4440  
QY 5157 ATTGGGTGTGATATTCCTGGCTGGTATTTGTGGGTCTCACTGAAGATGGCATCCCT 5216  
DB 4441 ATTGGGTGTGATATTCCTGGCTGGTATTTGTGGGTCTCACTGAAGATGGCATCCCT 4500  
QY 5217 TATCTGGCTGTGCTGATTTGTTTTTCCACTATTTTACTTGGGGTAACCTCCGCTCAGGAA 5276

DB 4501 TATCTGGCTGTGCTGCAATTTGTTTTTCCACTATTTACTTGGGGTAACCTCCGCTCAGGAA 4560  
QY 5277 CTGCATACCAATTTCTGCAGAGGAGAGTACAGTGCACCTCTGTAGCTATCTATCTTTACCT 5336  
DB 4561 CTGCATACCAATTTCTGCAGAGGAGAGTACAGTGCACCTCTGTAGCTATCTATCTTTACCT 4620  
QY 5337 ACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATACTGTAAAGCCCTTCTCTCCAGAGG 5396  
DB 4621 ACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATACTGTAAAGCCCTTCTCTCCAGAGG 4680  
QY 5397 TGGTGTGAGATCTGCTCTTAACTGTTTGAAGCAAAAAACAACGTTGGTCAAGTAC 5456  
DB 4681 TGGTGTGAGATCTGCTCTTAACTGTTTGAAGCAAAAAACAACGTTGGTCAAGTAC 4740  
QY 5457 CTTAGAAAAAAGAAATAGTTTGTATAGAGCTTCTGATGATATATAGCTGCTCTGATCAAA 5516  
DB 4741 CTTAGAAAAAAGAAATAGTTTGTATAGAGCTTCTGATGATATATAGCTGCTCTGATCAAA 4800  
QY 5517 GCTTCTCAATTTCCAGTGCCTCACCGTCTGCAGATGATGAGCGAAAAAGCATCTCTCTGTC 5576  
DB 4801 GCTTCTCAATTTCCAGTGCCTCACCGTCTGCAGATGATGAGCGAAAAAGCATCTCTCTGTC 4860  
QY 5577 CTTTCTGTGGGCTATATCTATGTTCTCAGAAATTTTGTCTGCCAGAAATTTGTGAACGGG 5636  
DB 4861 CTTTCTGTGGGCTATATCTATGTTCTCAGAAATTTTGTCTGCCAGAAATTTGTGAACGGG 4920  
QY 5637 GAAAGGTTGTGAGCTTGCATTTTTCAGCACTTCACTGTGGAGCGGAGTCTGCATTTTC 5696  
DB 4921 GAAAGGTTGTGAGCTTGCATTTTTCAGCACTTCACTGTGGAGCGGAGTCTGCATTTTC 4980  
QY 5697 CTAAAAATCAGAGAAATGCCAGTGTCTGTGTTGAAGTAAAGCCAGAGGCTGTGCCCTAT 5756  
DB 4981 CTAAAAATCAGAGAAATGCCAGTGTCTGTGTTGAAGTAAAGCCAGAGGCTGTGCCCTAT 5040  
QY 5757 CCAGCTCTTACTTGGATGAATATGGAAGAAACAGACCTGTGCTGGAAGGGGCAACCCC 5816  
DB 5041 CCAGCTCTTACTTGGATGAATATGGAAGAAACAGACCTGTGCTGGAAGGGGCAACCCC 5100  
QY 5817 CTTCAATTTCTGCTGAGCGGTATCGGAGCTTCCATTTGGTCTGCGCAACACCTGCAT 5876  
DB 5101 CTTCAATTTCTGCTGAGCGGTATCGGAGCTTCCATTTGGTCTGCGCAACACCTGCAT 5160  
QY 5877 ATAGAAGAGATTTGCTAGGAGCAAGAGACTAATCAGATGTTTATTTGGATTTCAACTGGCAG 5936  
DB 5161 ATAGAAGAGATTTGCTAGGAGCAAGAGACTAATCAGATGTTTATTTGGATTTCAACTGGCAG 5220  
QY 5937 TTACTGTGAGCTCAACTCTGCTCAAGACAATCAAAATGCAAGCAGTAGTAAGGCTG 5996  
DB 5221 TTACTGTGAGCTCAACTCTGCTCAAGACAATCAAAATGCAAGCAGTAGTAAGGCTG 5280  
QY 5997 ATTCAAAATTTAGGAAGAACTTCTGAGGCTGCGGAAAGTATTTGGAGGCTCTTTTGTCCA 6056  
DB 5281 ATTCAAAATTTAGGAAGAACTTCTGAGGCTGCGGAAAGTATTTGGAGGCTCTTTTGTCCA 5340  
QY 6057 TGTCCAGGCTCACTTACATCAATAAATAATTTCTTAATGAGATTTGCTTTCAATTAGCA 6116  
DB 5341 TGTCCAGGCTCACTTACATCAATAAATAATTTCTTAATGAGATTTGCTTTCAATTAGCA 5400  
QY 6117 AACATATGCTTTCAGAGGAAAAAGGACATAGATCAATCTGTTTATGCTGTAGTATTTCC 6176  
DB 5401 AACATATGCTTTCAGAGGAAAAAGGACATAGATCAATCTGTTTATGCTGTAGTATTTCC 5460  
QY 6177 AGGAATTTATTTCCCTTCAATAATTTGCTCAATTTTATTTTATTTTATTCATCCCTGTAGAT 6236  
DB 5461 AGGAATTTATTTCCCTTCAATAATTTGCTCAATTTTATTTTATTTTATTTTATTTTATTTT 5520  
QY 6237 GAAGTCACTCAAAACAGTTGTAGACATTTTATGTTGGTTAACTCTTCTGCAATTTTGT 6296  
DB 5521 GAAGTCACTCAAAACAGTTGTAGACATTTTATGTTGGTTAACTCTTCTGCAATTTTGT 5580  
QY 6297 ATTTGGTGTTTT 6308  
|||||



Db 5581 ATTGGTGT 5592

RESULT 4  
US-10-474-291-22  
; Sequence 22. Application US/10474291  
; Publication No. US20040132043A1  
; GENERAL INFORMATION:  
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.  
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.  
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.  
; APPLICANT: BURFORD, Neil; DING, Li  
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.  
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.A.  
; APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.  
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.  
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.  
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi  
; APPLICANT: REDDY, Roopa M.; SANTANWALA, Madhusudan M.  
; APPLICANT: TANG, Y. Tom; CHAWLA, Narinder K.  
; APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.  
; APPLICANT: XU, Yuming; YANG, Junming  
; APPLICANT: YAO, Monique G.; YUE, Henry  
; APPLICANT: ZEBARJADIAN, Yeganeh  
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH  
; FILE REFERENCE: PI-0417 USN  
; CURRENT APPLICATION NUMBER: US/10/474,291  
; CURRENT FILING DATE: 2003-10-06  
; PRIOR APPLICATION NUMBER: PCT/US02/11152  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: US 60/349,705  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: US 60/295,263  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/295,340  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/293,727  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 60/291,846  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/291,662  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: US 60/287,228  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/286,820  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/283,294  
; PRIOR FILING DATE: 2001-04-11  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 22  
; LENGTH: 7742  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 1351608CB1  
US-10-474-291-22.

Query Match 88.6%; Score 5592; DB 19; Length 7742;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 717 GGTACTGAGCGATGGAATCAGCGGGAGTACCCAGACCCCTCAGCGTCTGGCATCT 776  
Db 1 GGTACTGAGAGATGGAATCAGCGGGAGTACCCAGACCCCTCAGCGTCTGGCATCT 60

Qy 777 TGGTGGGATCAGCAAGTGTGATTTTATCTGCTTTCTTGCATCTTTGGCACAATTTGGTG 836  
Db 61 TGGTGGGATCAGCAAGTGTGATTTTATCTGCTTTCTTGCATCTTTGGCACAATTTGGTG 120

Qy 837 CCAGAAATTTACTTTTGTCTGAAATGGAACCCAGACTTTGGAAAGCAGGAGGAAGTGTACAA 896

Db 121 CCAGAAATTTACTTTTGTCTGAAATGGAACCCAGACTTTGGAAAGCAGGAGGAAGTGTACAA 180

Qy 897 ATGTCAATATTTCACTCCCACTGGAAATGGTACTTTATTTGGAGAAAGATCCAGATATTTGCTTTA 956

Db 181 ATGTCAATATTTCACTCCCACTGGAAATGGTACTTTATTTGGAGAAAGATCCAGATATTTGCTTTA 240

Qy 957 GAGAAATTTGAAGCACAGTGGAGCATTTTTCAGCTTTTGTGGAGGGTTTTCAAAAGTGGAGAG 1016

Db 241 GAGAAATTTGAAGCACAGTGGAGCATTTTTCAGCTTTTGTGGAGGGTTTTCAAAAGTGGAGAG 300

Qy 1017 ACAACCTATTTCTTGCAGGGATTTGCAATTTGATCCAAATGATGTAATCTCTGTATGACTGC 1076

Db 301 ACAACCTATTTCTTGCAGGGATTTGCAATTTGATCCAAATGATGTAATCTCTGTATGACTGC 360

Qy 1077 TTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGCACTACTTTCTACTGGAGGAGG 1136

Db 361 TTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGCACTACTTTCTACTGGAGGAGG 420

Qy 1137 TTCTGTGACTGTGGAGACACAGAGGCATGGAAAACCTGGCCCTTTTGTGTGTAATCATGAA 1196

Db 421 TTCTGTGACTGTGGAGACACAGAGGCATGGAAAACCTGGCCCTTTTGTGTGTAATCATGAA 480

Qy 1197 CCTGGAACAGCAGGTACTATATAAGAGAAATTCAGCTGTCGCTGGAATGGAAGAGCTAATT 1256

Db 481 CCTGGAACAGCAGGTACTATATAAGAGAAATTCAGCTGTCGCTGGAATGGAAGAGCTAATT 540

Qy 1257 GTCCAGGCCAGGAAATATTTCTTCACTGATATAAATATGTCGTAGAATGCACTATATGG 1316

Db 541 GTCCAGGCCAGGAAATATTTCTTCACTGATATAAATATGTCGTAGAATGCACTATATGG 600

Qy 1317 GAAGAGGAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAATGAAAGATACATAT 1376

Db 601 GAAGAGGAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAATGAAAGATACATAT 660

Qy 1377 TGTGTCTTTTCAATGATGAACACCATTCATATGACCACTCATATACAGCTACAAAGA 1436

Db 661 TGTGTCTTTTCAATGATGAACACCATTCATATGACCACTCATATACAGCTACAAAGA 720

Qy 1437 GCTCTTGACTGTGAGCTCGCAGAGCCAGTTTGCATACACTGCCATTCGACAAAGAGGT 1496

Db 721 GCTCTTGACTGTGAGCTCGCAGAGCCAGTTTGCATACACTGCCATTCGACAAAGAGGT 780

Qy 1497 CGTGGGCTGTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAGAAAGATATAAG 1556

Db 781 CGTGGGCTGTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAGAAAGATATAAG 840

Qy 1557 AGTCATTCAGAAATGCTCTCAACATCCACTTCATGTAGAGTATTTACACTCAGAGATT 1616

Db 841 AGTCATTCAGAAATGCTCTCAACATCCACTTCATGTAGAGTATTTACACTCAGAGATT 900

Qy 1617 ATGGCTCATCAGAAATTTGCTTGGCTCTTGGTTCCTGGATGAAACAAATTTATGAGCTAT 1676

Db 901 ATGGCTCATCAGAAATTTGCTTGGCTCTTGGTTCCTGGATGAAACAAATTTATGAGCTAT 960

Qy 1677 TCAAGTGAATTTAGGACAGATCTTTTGGCAAGCATGCTCTTAGAGAGAAACCTGACTCGGAG 1736

Db 961 TCAAGTGAATTTAGGACAGATCTTTTGGCAAGCATGCTCTTAGAGAGAAACCTGACTCGGAG 1020

Qy 1737 AATCCCTGCTCATAGCAGGTAAATGCTTTGGGATGCAAGCTTTATAAGTGGCCCT 1796

Db 1021 AATCCCTGCTCATAGCAGGTAAATGCTTTGGGATGCAAGCTTTATAAGTGGCCCT 1080

Qy 1797 AAGATCCTTCATGAATTTGATCTTCAGCAGTCTTTTATGAGAGATGGAATACAAAAAATC 1856

Db 1081 AAGATCCTTCATGAATTTGATCTTCAGCAGTCTTTTATGAGAGATGGAATACAAAAAATC 1140

Qy 1857 TTTGCTATGGAATTTGTGAAGTATTTATAAACAACCTGCAAGAAAGATATATCAGTGATGAT 1916

Db 1141 TTTGCTATGGAATTTGTGAAGTATTTATAAACAACCTGCAAGAAAGATATATCAGTGATGAT 1200

Qy 1917 CATGACAGAGATATCTCTATTAACCTGCACCTTCAGTTCAGATGTTTACTGTTCTTACTCTG 1976

Db 1201 CATGACAGAGTATCTCTATAACTGCACCTTTTCAGTTTCAGATGTTTACTGTTCTTACTCTG 1260  
Qy 1977 GCTCGACATCTTTATTTGAAGAGCAGAAATGTTATCTCTGTCTCATTTACTGAACTCTGCTAGAA 2036  
Db 1261 GCTCGACATCTTTATTTGAAGAGCAGAAATGTTATCTCTGTCTCATTTACTGAACTCTGCTAGAA 1320  
Qy 2037 GTTTTACCTGAGTACTTGGACAGGAAACAATAAATCAACTCTCCAGGGTTATAGCCAGGAC 2096  
Db 1321 GTTTTACCTGAGTACTTGGACAGGAAACAATAAATCAACTCTCCAGGGTTATAGCCAGGAC 1380  
Qy 2097 AAATTTGGGAAGAGTATATGTCAGTAAATATGTCAGCTTAAAGTATATCTCTGATCAGCAAAACC 2156  
Db 1381 AAATTTGGGAAGAGTATATGTCAGTAAATATGTCAGCTTAAAGTATATCTCTGATCAGCAAAACC 1440  
Qy 2157 ACAATATGGACAGAAAGATTAAAGATGAGTTCTTCTGAAAGTTTTCGATCTTTTTCGAAG 2216  
Db 1441 ACAATATGGACAGAAAGATTAAAGATGAGTTCTTCTGAAAGTTTTCGATCTTTTTCGAAG 1500  
Qy 2217 ATTCTTACCTGTATGTCAGGGAATGAAAGAAATCCGAAGACAGGTTGGGCAACACATTGAA 2276  
Db 1501 ATTCTTACCTGTATGTCAGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTGAA 1560  
Qy 2277 GTGATCTCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAATATTTTACTC 2336  
Db 1561 GTGGATCTCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAATATTTTACTC 1620  
Qy 2337 ATGTTTCAAGAGTGTGTGCTTGTGATGAGAACTCTTACTTGTGGCTTATTAAGAATGT 2396  
Db 1621 ATGTTTCAAGAGTGTGTGCTTGTGATGAGAACTCTTACTTGTGGCTTATTAAGAATGT 1680  
Qy 2397 CACAAGCTGTATGTCAGGTCAGTACCACTTTTCATATCTAGTAGCAAGCAGTAGTACAA 2456  
Db 1681 CACAAGCTGTATGTCAGGTCAGTACCACTTTTCATATCTAGTAGCAAGCAGTAGTACAA 1740  
Qy 2457 TGTGTGGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTTGTAAGCATA 2516  
Db 1741 TGTGTGGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTTGTAAGCATA 1800  
Qy 2517 CATCTGCCATCTCTAGGACCTTCTGCTGTCTTCACTGATACGTTTAAAGCAGGCTGGGTCT 2576  
Db 1801 CATCTGCCATCTCTAGGACCTTCTGCTGTCTTCACTGATACGTTTAAAGCAGGCTGGGTCT 1860  
Qy 2577 GTTTCGAAGCTCATGAAATTTGTCTTTTGAAGACTTTTCAAGTAGAGGTACTAGTGGAA 2636  
Db 1861 GTTTCGAAGCTCATGAAATTTGTCTTTTGAAGACTTTTCAAGTAGAGGTACTAGTGGAA 1920  
Qy 2637 TATCTCTTACGTGTCTGTGTGTGTCGCCAGGTTGTGTGAGATGTGGCGAAGAAAT 2696  
Db 1921 TATCTCTTACGTGTGTGTGTGTCGCCAGGTTGTGTGAGATGTGGCGAAGAAAT 1980  
Qy 2697 GGACTGTCTCTTATTTAGCCAGGTGTTTATACCAAGATGTTTAAAGTCAGAGAAATG 2756  
Db 1981 GGACTGTCTCTTATTTAGCCAGGTGTTTATTTACCAAGATGTTTAAAGTCAGAGAAATG 2040  
Qy 2757 TATGATAAAGATATCATCATGCTTCAGATTGTGTGATCTTTAATGATCCCAATTAAGTTC 2816  
Db 2041 TATGATAAAGATATCATCATGCTTCAGATTGTGTGATCTTTAATGATCCCAATTAAGTTC 2100  
Qy 2817 TTGTTACTGCTTCTCAGAGGTATGAACTTGGCCGAGGCTTTTAAACAAGACCAATCTACA 2876  
Db 2101 TTGTTACTGCTTCTCAGAGGTATGAACTTGGCCGAGGCTTTTAAACAAGACCAATCTACA 2160  
Qy 2877 AAAGACCCAGGATTTGATTTAAACAATATATACATTAATAGAAGAAATGCTTCAGGTCTCT 2936  
Db 2161 AAAGACCCAGGATTTGATTTAAACAATATATACATTAATAGAAGAAATGCTTCAGGTCTCT 2220  
Qy 2937 ATCTATATTTGCGGTGAGCGTTATGTACTGAGTGGGAAATGTGACCAAGAGAGGTC 2996  
Db 2221 ATCTATATTTGCGGTGAGCGTTATGTACTGAGTGGGAAATGTGACCAAGAGAGGTC 2280  
Qy 2997 ACAATGAGAGAAATCATCTGCTTGTGCAATTTGAACCCATGCCACACAGTCCCATGCC 3056  
Db 2281 ACAATGAGAGAAATCATCTGCTTGTGCAATTTGAACCCATGCCACACAGTCCCATGCC 2340

Qy 3057 AAAAATTTTACCTGAGAAATGAAAATAATGAAACTCGCTTTAGAGAAATGTCTATAAACCAAGTG 3116  
Db 2341 AAAAATTTTACCTGAGAAATGAAAATAATGAAACTCGCTTTAGAGAAATGTCTATAAACCAAGTG 2400  
Qy 3117 GCCACATTTAAGAAAACAGAGTGTATCAGGGCAGAGTGTATATGAACTAAAAAGATGAATCA 3176  
Db 2401 GCCACATTTAAGAAAACAGAGTGTATCAGGGCAGAGTGTATATGAACTAAAAAGATGAATCA 2460  
Qy 3177 CTGAAGACTTCAATATATGTTACTTTTATCTTATCTTCCAAACCCAGCATAGCAAGGCTGAA 3236  
Db 2461 CTGAAGACTTCAATATATGTTACTTTTATCTTATCTTCCAAACCCAGCATAGCAAGGCTGAA 2520  
Qy 3237 CATATGCAAGAAAACAGAGAAAACAAGAAAACAAGATGAAGCAATTTGCCGCCACCAACA 3296  
Db 2521 CATATGCAAGAAAACAGAGAAAACAAGAAAACAAGATGAAGCAATTTGCCGCCACCAACA 2580  
Qy 3297 CCTCCTGAATTTCTGCCCTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTGATATCATG 3356  
Db 2581 CCTCCTGAATTTCTGCCCTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTGATATCATG 2640  
Qy 3357 ATGTACATTTCTCAGGACCGTATTTTGGGGGCAATAGACACAGATTTCTAACTTTGGGACC 3416  
Db 2641 ATGTACATTTCTCAGGACCGTATTTTGGGGGCAATAGACACAGATTTCTAACTTTGGGACC 2700  
Qy 3417 GAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGCAATTTGGGTTTCTAGAAAGAGAACAA 3476  
Db 2701 GAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGCAATTTGGGTTTCTAGAAAGAGAACAA 2760  
Qy 3477 CAGCTTCAAAAAGCTCTCAGAAAGAAAGTAACTTTTGAATTTTATCATTAAGGCTTCAAGA 3536  
Db 2761 CAGCTTCAAAAAGCTCTCAGAAAGAAAGTAACTTTTGAATTTTATCATTAAGGCTTCAAGA 2820  
Qy 3537 TTGGGAAGTTTTCAGCCATGAAATATACAAATGCTTTTGGAAAACCTCAAGGAAATTTCCCCAG 3596  
Db 2821 TTGGGAAGTTTTCAGCCATGAAATATACAAATGCTTTTGGAAAACCTCAAGGAAATTTCCCCAG 2880  
Qy 3597 TTAGAAGCCAGAAAGGACATGATAACGTTGGATATCTTCAGATGTTTGAACAGTGAAGCGA 3656  
Db 2881 TTAGAAGCCAGAAAGGACATGATAACGTTGGATATCTTCAGATGTTTGAACAGTGAAGCGA 2940  
Qy 3657 TTAAGAGAAAATCTTTGTTTAAATTTAGCAACCAATCAGGATCGGAATCTTATTAAGAAAT 3716  
Db 2941 TTAAGAGAAAATCTTTGTTTAAATTTAGCAACCAATCAGGATCGGAATCTTATTAAGAAAT 3000  
Qy 3717 GATGAGATTACTCATGATAAAGAAAAGCAGAAACGAAAAGAAAGCTGGAAGCTGCTAGG 3776  
Db 3001 GATGAGATTACTCATGATAAAGAAAAGCAGAAACGAAAAGAAAGCTGGAAGCTGCTAGG 3060  
Qy 3777 CTACATGCCAGAAAGATCATGGCTCAGATGCTGCTTTACAGAAAACCTTCATTGAAAAT 3836  
Db 3061 CTACATGCCAGAAAGATCATGGCTCAGATGCTGCTTTACAGAAAACCTTCATTGAAAAT 3120  
Qy 3837 CATAACTCATGTATGACAAATACATCAAGAAATGCTTGGGAAAGAAAGATTCATTATGAG 3896  
Db 3121 CATAACTCATGTATGACAAATACATCAAGAAATGCTTGGGAAAGAAAGATTCATTATGAG 3180  
Qy 3897 GAAGAGACCCACAGCCAGTCACTACTCTAGAAATTTGGGTCTTAAACGGGT 3956  
Db 3181 GAAGAGACCCACAGCCAGTCACTACTCTAGAAATTTGGGTCTTAAACGGGT 3240  
Qy 3957 CCATCTGTTACTGAAAAGGAGGTGCTGACGTGCACTTTTCCCAAGAAAGAAACAGGAGTG 4016  
Db 3241 CCATCTGTTACTGAAAAGGAGGTGCTGACGTGCACTTTTCCCAAGAAAGAAACAGGAGTG 3300  
Qy 4017 AAAATAGAAAATAATGCCATGTTATTCGCCCTGTGTCCAGAAATCTACTGCTTAAACC 4076  
Db 3301 AAAATAGAAAATAATGCCATGTTATTCGCCCTGTGTCCAGAAATCTACTGCTTAAACC 3360  
Qy 4077 CAGCACAGGGGAAAACCCATGAACTCTCAGAGAGAGCCCTAGACCCACTTTTCAATGAT 4136  
Db 3361 CAGCACAGGGGAAAACCCATGAACTCTCAGAGAGAGCCCTAGACCCACTTTTCAATGAT 3420

4137 CCAGACTTGGCATATGGAACTTATACAGGAAGCTGTGGTCACTGAATGCAACGAGTGGC 4196  
4217 CCAGACTTGGCATATGGAACTTATACAGGAAGCTGTGGTCACTGAATGCAACGAGTGGC 4256  
4257 TGGCAGAACTATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTGACCTT 3540  
4257 TTTGACTTGGAAAGTGGAGAAATATCTTTGGCCCTCTTTGCAAAATCTCTGTGCAATATCTGTG 4316  
4317 ATCCCATTTATCTTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTGACCTT 3600  
4317 ATCCCATTTATCTTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTGACCTT 4376  
4376 ATCCCATTTATCTTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTGACCTT 3660  
4377 CAACCTTTTGACCTTGGCAGCGTGGATACAGACTGTCTTGGCCAGAAATATCAGGTTATAT 4436  
4377 CAACCTTTTGACCTTGGCAGCGTGGATACAGACTGTCTTGGCCAGAAATATCAGGTTATAT 3720  
4437 ATAGAGCATGCTAAAGGAGAAACCCAAATTCCTATTTTCTTTAAATCAAGAAATGGAGAT 4496  
4497 ATAGAGCATGCTAAAGGAGAAACCCAAATTCCTATTTTCTTTAAATCAAGAAATGGAGAT 3780  
4497 TCTACTTTGGAGTTCATTTCCATCTGAGTGTGCGTTTGGCGTTGAGTCTTCGATTAATATTTCA 4556  
4557 ATAGCATCAAGAAATGGTATTTCTTTTGGCAAAATTTATAGAAATGGATTTGAAA 4616  
4617 ATAGCATCAAGAAATGGTATTTCTTTTGGCAAAATTTATAGAAATGGATTTGAAA 3900  
4617 GTGCCACCTGATGAAAGGAGTCTCGAGTCCCCTGCTGACCTGGAGCACCTGGCCTTTC 4676  
4677 GTGCCACCTGATGAAAGGAGTCTCGAGTCCCCTGCTGACCTGGAGCACCTGGCCTTTC 3960  
4677 ACTATCCAGGCAATTTGAAATCTATTTGGAGATGAAAGAAACCTCTGTGTTGGAGCACTT 4736  
4737 ACTATCCAGGCAATTTGAAATCTATTTGGAGATGAAAGAAACCTCTGTGTTGGAGCACTT 4020  
4737 CAATAATAGGCAATTAAGTCTGAAAGCAATTAAGTCTGAGTTGCAAGTGGACAGAGGAT 4796  
4801 CAATAATAGGCAATTAAGTCTGAAAGCAATTAAGTCTGAGTTGCAAGTGGACAGAGGAT 4080  
4857 ACCTGTCTCAGGTCCTGATACAGAAACATCTGGTCTGCTATCAGTGTCTCTCT 4856  
4857 ACCTGTCTCAGGTCCTGATACAGAAACATCTGGTCTGCTATCAGTGTCTCTCT 4140  
4857 AACATAAATCAGAAAGATACACCATGCTCTGTCTATAGATCTGTCTTCAAGTGTGGTG 4916  
4917 AACATAAATCAGAAAGATACACCATGCTCTGTCTATAGATCTGTCTTCAAGTGTGGTG 4200  
4917 GGTGCTGTGTAGCAATCCCATCTTGTATTTGGAGTACCCCTGTGATCTGAGCTTCT 4976  
4977 TCAGTTAGTCTTCTTATAACCACTTTATCTCTTCCATTTGATCAACATGGCACACATG 4260  
4977 TCAGTTAGTCTTCTTATAACCACTTTATCTCTTCCATTTGATCAACATGGCACACATG 5036  
4977 TCAGTTAGTCTTCTTATAACCACTTTATCTCTTCCATTTGATCAACATGGCACACATG 4320  
5037 CTCAGATATCTTACAGTATAGACACAGGCTTACCCCTTGTCTGAGTTCAAGAAAGAT 5096  
5097 CTCAGATATCTTACAGTATAGACACAGGCTTACCCCTTGTCTGAGTTCAAGAAAGAT 4380  
5097 GAAAGGCTCATTCGCACTCTCTTTCTTTGAGAAATTTCTCAATATACAGTGGCTCC 5156  
5157 GAAAGGCTCATTCGCACTCTCTTTCTTTGAGAAATTTCTCAATATACAGTGGCTCC 4440  
5157 ATTTGGGTGTATATCTTGGCTGTATTTGAGGCTCTCACTGAAGAAATGGCATCACCCCT 5216  
5217 ATTTGGGTGTATATCTTGGCTGTATTTGAGGCTCTCACTGAAGAAATGGCATCACCCCT 4500  
5217 TATCTTCTGCTGTGCTATTTTTCATATTTTCTGAGGTTAACTCCGCTGAGGAA 5276

4501 TATCTTCTGCTGTGCTATTTTTCATATTTTCTTGGGTAATCTCGCTCAGGAA 4560  
5277 CTGCAATACCAATTTCTGCAAGAGAGATGACAGTGCATCTGTAGCTATCTATCTTTACCT 5336  
5277 CTGCAATACCAATTTCTGCAAGAGAGATGACAGTGCATCTGTAGCTATCTATCTTTACCT 4620  
5337 ACATAATTTGTTCTGCTCTCTTCCAGGAATATTTGGGATCTGTAAAGGCTTGTCTCAGAGG 5396  
5397 ACATAATTTGTTCTGCTCTCTTCCAGGAATATTTGGGATCTGTAAAGGCTTGTCTCAGAGG 4680  
5397 TGGTGTGAGATCTTCTGCTCTTCTTAACTGTTTGAAGCAAAAAACAACCGTGTGAGTAC 5456  
5457 TGGTGTGAGATCTTCTGCTCTTCTTAACTGTTTGAAGCAAAAAACAACCGTGTGAGTAC 4740  
5457 CCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTGTATGATCTATAGCTGCTCTCTGATCAA 5516  
5457 CCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTGTATGATCTATAGCTGCTCTCTGATCAA 4800  
5517 GCTTCTCATTTTCTGAGTGGCCAGCTGTGAGATGATGAGCGAAAGATCTCTGCTCTGCTG 5576  
5577 GCTTCTCATTTTCTGAGTGGCCAGCTGTGAGATGATGAGCGAAAGATCTCTGCTCTGCTG 4860  
5577 CTTTTCTGTTGGGCTTATCTATGTTCTCAGAAATTTGCTGCGCAAGAAATTTGTAACGG 5636  
5637 GAAAGGTTTGGAGCTTGCATTTTTCACGCACTTCTGAGAGCGGAGTCTGCAATTTTC 4920  
5637 GAAAGGTTTGGAGCTTGCATTTTTCACGCACTTCTGAGAGCGGAGTCTGCAATTTTC 5696  
5697 GAAAGGTTTGGAGCTTGCATTTTTCACGCACTTCTGAGAGCGGAGTCTGCAATTTTC 4980  
5697 GTAAAAATCAGAAATGCGAGTGTCTGTTTGAAGTAAAGCAGAGGCTGTGCTAT 5756  
5757 GTAAAAATCAGAAATGCGAGTGTCTGTTTGAAGTAAAGCAGAGGCTGTGCTAT 5040  
5757 CCAGCTCTTCTTGGATGAAATATGGAACACAGACCTTGGCTCAAGAGGGGCAACCC 5816  
5817 CCAGCTCTTCTTGGATGAAATATGGAACACAGACCTTGGCTCAAGAGGGGCAACCC 5100  
5817 CTTCAATTTATCTGAGGCTTATCGGAGCTTCAATTTGCTGCGCAACCACTGCTAT 5876  
5877 CTTCAATTTATCTGAGGCTTATCGGAGCTTCAATTTGCTGCGCAACCACTGCTAT 5160  
5877 ATAGAGAGATTTGCTAGGCAAGAGACTAATCAAGTGTATTTTGAATTTCACTGGCAG 5936  
5937 ATAGAGAGATTTGCTAGGCAAGAGACTAATCAAGTGTATTTTGAATTTCACTGGCAG 5220  
5937 TTAATGAGCTTCACTGCTGCTCAAGCAATCAATAATGAGCAAGTGTAAAGGCTG 5996  
5997 TTAATGAGCTTCACTGCTGCTCAAGCAATCAATAATGAGCAAGTGTAAAGGCTG 5280  
5997 ATTCAAAATTTGAAATTTTCTGAGGCTTGGGAAAGTATTTGAGGCTTCTTCTGCTCA 6056  
5997 ATTCAAAATTTGAAATTTTCTGAGGCTTGGGAAAGTATTTGAGGCTTCTTCTGCTCA 5340  
6057 TGTCCAGGTTCACTTACATCAATTAATTTCTTAAATGAGTATTTGCTTTCAATAGCA 6116  
6117 TGTCCAGGTTCACTTACATCAATTAATTTCTTAAATGAGTATTTGCTTTCAATAGCA 5400  
6117 AACATATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTGTAGTATTTCC 6176  
6177 AACATATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTGTAGTATTTCC 5460  
6177 AGGAATTTATTTCCCTTCTCATATTTTCTCATTTTCTCATTTTCTCATTTTCTCATTTTCT 6236  
6237 AGGAATTTATTTCCCTTCTCATATTTTCTCATTTTCTCATTTTCTCATTTTCTCATTTTCT 5520  
6237 GAACTCAGTCAAAACAGTGTAGACATTTTATGTTGTTGTTAACTCTTCTGCAATTTTGT 6296  
6297 GAACTCAGTCAAAACAGTGTAGACATTTTATGTTGTTGTTAACTCTTCTGCAATTTTGT 5580  
6297 ATTTGGGTGTTT 6308  
|||||

Db 5581 ATTGGTGTGTTT 5592

RESULT 5

US-10-357-819-9

; Sequence 9, Application US/10357819

; Publication No. US20040259774A1

; GENERAL INFORMATION:

; APPLICANT: Alvarez, Enrique

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Gangolli, Esha A.

; APPLICANT: Gerlach, Valerie L.

; APPLICANT: Gorman, Linda

; APPLICANT: Guo, Xiaojia

; APPLICANT: Ji, Weizhen

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Li, Li

; APPLICANT: Miller, Charles E.

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Pattursajan, Meera

; APPLICANT: Rastelli, Luca

; APPLICANT: Rieger, Daniel K.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Zhong, Mei

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

; FILE REFERENCE: 21402-538A

; CURRENT APPLICATION NUMBER: US/10/357,819

; PRIOR FILING DATE: 2003-02-03

; PRIOR APPLICATION NUMBER: 09/520,781

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 09/584,411

; PRIOR FILING DATE: 2000-05-31

; PRIOR APPLICATION NUMBER: 09/783,436

; PRIOR FILING DATE: 2001-02-14

; PRIOR APPLICATION NUMBER: 10/085,198

; PRIOR FILING DATE: 2002-02-25

; PRIOR APPLICATION NUMBER: 60/353,301

; PRIOR FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 60/355,099

; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: 60/356,424

; PRIOR FILING DATE: 2002-02-12

; PRIOR APPLICATION NUMBER: 60/358,239

; PRIOR FILING DATE: 2002-02-20

; PRIOR APPLICATION NUMBER: 60/358,608

; PRIOR FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: 60/359,367

; PRIOR FILING DATE: 2002-02-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 142

; SOFTWARE: CuraSeqList version 0.1

; SEQ ID NO 9

; LENGTH: 5466

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (66)..(5312)

US-10-357-819-9

Query Match 86.7%; Score 5466; DB 20; Length 5466;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 631 CGGCACAGTTTCGCTGCTCTGCGCGGGGTCGCAACTGCGAGGCTCAGTTTCCC 690

Db 1 CGGCACAGTTTCGCTGCTCTGCGCGGGGTCGCAACTGCGAGGCTCAGTTTCCC 60

QY 691 TCAAGATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGGAGTTAC 750

Db 61 TCAAGATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGGAGTTAC 120

QY 751 CCCAGACCCCTCAGCGCTCGCATCTTGTGGGATCAGCAAGTTGATTTTTTATCTGCTT 810

Db 121 CCCAGACCCCTCAGCGCTCGCATCTTGTGGGATCAGCAAGTTGATTTTTTATCTGCTT 180

QY 811 TCTTGCAATCAATTTGGGCACAAATGTGTGCCAGAAAATTTACTTTGTGGAATGGACCCAGACT 870

Db 181 TCTTGCAATCAATTTGGGCACAAATGTGTGCCAGAAAATTTACTTTGTGGAATGGACCCAGACT 240

QY 871 TGGAAAAGCAGGAGGAAGTGTACAAATGTCAATTAATTCATCTCCAATGGAATGGTACTTAT 930

Db 241 TGGAAAAGCAGGAGGAAGTGTACAAATGTCAATTAATTCATCTCCAATGGAATGGTACTTAT 300

QY 931 TTGGAGAAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCACAAGTGGAGCAATTTTCAGCTTT 990

Db 301 TTGGAGAAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCACAAGTGGAGCAATTTTCAGCTTT 360

QY 991 GTGGAGGGTTTTTCAAAGTGGAGAGACAACCTATTCTTTCAGGGATTTGTGCAATTTGATC 1050

Db 361 GTGGAGGGTTTTTCAAAGTGGAGAGACAACCTATTCTTTCAGGGATTTGTGCAATTTGATC 420

QY 1051 CAACATGTACTCTCTGTATGGACTGCTTCCAGGACAGTGTTCATATAAAATCATCGTTACA 1110

Db 421 CAACATGTACTCTCTGTATGGACTGCTTCCAGGACAGTGTTCATATAAAATCATCGTTACA 480

QY 1111 AGATGCATCTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGCATGGAAAA 1170

Db 481 AGATGCATCTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGCATGGAAAA 540

QY 1171 CTGGCCCTTTTTTGTGTAAATCATGAACCTGGAAGCAGAGTACTATATAAGAGAAATTCAC 1230

Db 541 CTGGCCCTTTTTTGTGTAAATCATGAACCTGGAAGCAGAGTACTATATAAGAGAAATTCAC 600

QY 1231 GCTGTCCGTTGAATGAAGAGGTAATTTGTCGAAGCCAGGAAATATTTCTTCAGTGATAA 1290

Db 601 GCTGTCCGTTGAATGAAGAGGTAATTTGTCGAAGCCAGGAAATATTTCTTCAGTGATAA 660

QY 1291 AATATGTCGTAGAAATGCACTATATGGAAGAGGAAAGAAAGAACTGCTCTCTCAATCATATG 1410

Db 721 AAGGAGAAATGCACTATATGGAAGAGGAAAGAAAGAACTGCTCTCTCAATCATATG 780

QY 1411 ACCAGCTCATATACAGCTTACAAAGAGCTCTTGAGCTGTGAGCTGCGAGGCCCAGTTGC 1470

Db 781 ACCAGCTCATATACAGCTTACAAAGAGCTCTTGAGCTGTGAGCTGCGAGGCCCAGTTGC 840

QY 1471 ATACCCTGCCATTTGACAAAGAGGCTGTCGGGCTGTTTAAAGCGGAGCTTATGCTGCTT 1530

Db 841 ATACCCTGCCATTTGACAAAGAGGCTGTCGGGCTGTTTAAAGCGGAGCTTATGCTGCTT 900

QY 1531 GCCAGGAAGCAAGGAAGATATAAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACTTC 1590

Db 901 GCCAGGAAGCAAGGAAGATATAAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACTTC 960

QY 1591 ATGTAGAAGTATTAACCTCAGAGATTAATGGCTCATCAGAAAATTTGCTTGGCTCTTGCTT 1650

Db 961 ATGTAGAAGTATTAACCTCAGAGATTAATGGCTCATCAGAAAATTTGCTTGGCTCTTGCTT 1020

QY 1651 CCTGGATGACAAAATTTATGAGCTATTCAAGTGACTTTTAGGCAGATCTTTTGGCCAAGCAT 1710

Db 1021 CCTGGATGACAAAATTTATGAGCTATTCAAGTGACTTTTAGGCAGATCTTTTGGCCAAGCAT 1080

QY 1711 GCCTTAGAAGAAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTTAATGCTTTGGG 1770

Db 1081 GCCTTAGAAGAAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTTAATGCTTTGGG 1140

QY 1771 ATGCAAGCTTTTAAAGGTGCGGTAAGATCCCTTCATGAATTTGATCTTCAGCAGTTTTT 1830

Db 1141 ATGCAAGCTTTTAAAGGTGCGGTAAGATCCCTTCATGAATTTGATCTTCAGCAGTTTTT 1200

1831 QY TTATGAGATGAATACAAAAAATCTTTGCTATGGAATTTGTGAAGTATTATAACAAAC 1890  
1201 Db TTATGAGATGGAATACAAAAAATCTTTGCTATGGAATTTGTGAAGTATTATAACAAAC 1260  
1891 QY TGCAGAAAGATATATACGTGATGATCATGACAGAGATATCTATAAATGCACTTTTCAG 1950  
1261 Db TGCAGAAAGATATATACGTGATGATCATGACAGAGATATCTATAAATGCACTTTTCAG 1320  
1951 QY TTACAGATTTTACTGTTCTTACTCTGGCTCGACATCTTATGAGAGAGAGATGTTATCT 2010  
1321 Db TTACAGATTTTACTGTTCTTACTCTGGCTCGACATCTTATGAGAGAGAGATGTTATCT 1380  
2011 QY CTGTCAATTAAGTAACTCTGCTAGAGATTTTACCTGAGTACTTGGACAGAGAAATTAAT 2070  
1381 Db CTGTCAATTAAGTAACTCTGCTAGAGATTTTACCTGAGTACTTGGACAGAGAAATTAAT 1440  
2071 QY TCAACTTCCAGGGTTATAGCCAGGACAAATTTGGAGAGATATATGAGTAAATATGTGACC 2130  
1441 Db TCAACTTCCAGGGTTATAGCCAGGACAAATTTGGAGAGATATATGAGTAAATATGTGACC 1500  
2131 QY TAAAGTATATCTGATCAGCAAAACCAATATGAGAGAGATTTAAGAAATGCAAGTTCC 2190  
1501 Db TAAAGTATATCTGATCAGCAAAACCAATATGAGAGAGATTTAAGAAATGCAAGTTCC 1560  
2191 QY TTGAAGGTTTTCGATCTTTTGAAGATTTTACCTGATGAGAGAGATTTAAGAAATGCAAGTTCC 2250  
1561 Db TTGAAGGTTTTCGATCTTTTGAAGATTTTACCTGATGAGAGAGATTTAAGAAATGCAAGTTCC 1620  
2251 QY GAAGACAGGTTGGGCAACATTTGAAGTGGATCTGATTTGGAGGCTGCAATTTGCTATAC 2310  
1621 Db GAAGACAGGTTGGGCAACATTTGAAGTGGATCTGATTTGGAGGCTGCAATTTGCTATAC 1680  
2311 QY AGATGCAATTTGAAGATATTTTACTCATGTTTCAAGAGTGGTGTCTTGTGATGAGAAC 2370  
1681 Db AGATGCAATTTGAAGATATTTTACTCATGTTTCAAGAGTGGTGTCTTGTGATGAGAAC 1740  
2371 QY TCTTACTTGTGCTTATTAAGATGTCACAAAGCTGTGATGAGTGGTGTGATGAGAAC 2430  
1741 Db TCTTACTTGTGCTTATTAAGATGTCACAAAGCTGTGATGAGTGGTGTGATGAGAAC 1800  
2431 QY TATCTAGTAGCAAGACAGTAGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCTTACA 2490  
1801 Db TATCTAGTAGCAAGACAGTAGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCTTACA 1860  
2491 QY GAGTATCTGAGGATCTTGTAGACATACATCTGCCACTCTCTAGGACCTTGTGCTTTC 2550  
1861 Db GAGTATCTGAGGATCTTGTAGACATACATCTGCCACTCTCTAGGACCTTGTGCTTTC 1920  
2551 QY ATGTAGGTTTAAAGAGGCTGGGTGCTGTTTCAAGACTGACATGAATTTGTCTTTTGAGG 2610  
1921 Db ATGTAGGTTTAAAGAGGCTGGGTGCTGTTTCAAGACTGACATGAATTTGTCTTTTGAGG 1980  
2611 QY ACTTTCAAGTAGAGTACTAGTGGAAATATCTTTTACGTTGCTGTTGTTGGTGGCCAGG 2670  
1981 Db ACTTTCAAGTAGAGTACTAGTGGAAATATCTTTTACGTTGCTGTTGTTGGTGGCCAGG 2040  
2671 QY TTGTGCTGAGATGTGGCAAGAAATGGAATGCTCTTATAGCAGAGTGTATTTATACC 2730  
2041 Db TTGTGCTGAGATGTGGCAAGAAATGGAATGCTCTTATAGCAGAGTGTATTTATACC 2100  
2731 QY AAGATGTTAAGTGACAGAGAAATGATGATAAAGATATCATGCTTCAGATTTGGT 2790  
2101 Db AAGATGTTAAGTGACAGAGAAATGATGATAAAGATATCATGCTTCAGATTTGGT 2160  
2791 QY CATCTTTAATGATCCCAATAGTTCTTTGTTACTGTGATCTTACAGAGTATGAACTTGGG 2850  
2161 Db CATCTTTAATGATCCCAATAGTTCTTTGTTACTGTGATCTTACAGAGTATGAACTTGGG 2220  
2851 QY AGGCTTTTAAAGACCAATATCTACAAAGACAGAGTTTGAATTAACAAATATATACAC 2910  
2221 Db AGGCTTTTAAAGACCAATATCTACAAAGACAGAGTTTGAATTAACAAATATATACAC 2280  
2911 QY TAATAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGTGAGCGTTATGTACCTGGAG 2970

2281 Db TAATAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGTGAGCGTTATGTACCTGGAG 2340  
2971 QY TGGGAAATGTGACCAAGAGAGAGGTCACAATGAGAGAAATCAATTCATCTTGTGCAATG 3030  
2341 Db TGGGAAATGTGACCAAGAGAGAGGTCACAATGAGAGAAATCAATTCATCTTGTGCAATG 2400  
3031 QY AACCATGCGCACACAGTGCATTTGCCAAAAATTTACCTGAGAAATGAATAATGAACATG 3090  
2401 Db AACCATGCGCACACAGTGCATTTGCCAAAAATTTACCTGAGAAATGAATAATGAACATG 2460  
3091 QY GCTTAGAGAAATGCTATAAACCAAGTGGCCACATTTTAAGAAACCAAGTGTATCAGGCCATG 3150  
2461 Db GCTTAGAGAAATGCTATAAACCAAGTGGCCACATTTTAAGAAACCAAGTGTATCAGGCCATG 2520  
3151 QY GAGTTTATGAACTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACT 3210  
2521 Db GAGTTTATGAACTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACT 2580  
3211 QY CCAAAACCCAGCATAGCAAGGCTGAACATATGACAGAGAAAGAGAGAAACAAAGAAACA 3270  
2581 Db CCAAAACCCAGCATAGCAAGGCTGAACATATGACAGAGAAAGAGAGAAACAAAGAAACA 2640  
3271 QY AAGATGAAGCATTTGCCGCCACCACTCTGTAATTTCTGCCCTGCTTTTCAGCAAGTGA 3330  
2641 Db AAGATGAAGCATTTGCCGCCACCACTCTGTAATTTCTGCCCTGCTTTTCAGCAAGTGA 2700  
3331 QY TTAACTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTAGCGGGCAA 3390  
2701 Db TTAACTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTAGCGGGCAA 2760  
3391 QY TAGACACAGATTTCTAACTTTGTGGAACGAGGATGCTCCAAATGCTTTTCAATTTCTGG 3450  
2761 Db TAGACACAGATTTCTAACTTTGTGGAACGAGGATGCTCCAAATGCTTTTCAATTTCTGG 2820  
3451 QY CATTTGGGTTTACTAGACAGAGCAACAGCTTCAAAAAGCTCCTCAAGAGAGTAAACAT 3510  
2821 Db CATTTGGGTTTACTAGACAGAGCAACAGCTTCAAAAAGCTCCTCAAGAGAGTAAACAT 2880  
3511 QY TTGACTTTTATCATAAAGGCTTCAAGATTTGGGAAGTTCCAGCATGAATATACAAATGCTTT 3570  
2881 Db TTGACTTTTATCATAAAGGCTTCAAGATTTGGGAAGTTCCAGCATGAATATACAAATGCTTT 2940  
3571 QY TGGAAAACTCAAGGAAATTTCCAGTTAGAGGCCAGAGGACATGATTAACGTGATAC 3630  
2941 Db TGGAAAACTCAAGGAAATTTCCAGTTAGAGGCCAGAGGACATGATTAACGTGATAC 3000  
3631 QY TTCAGATGTTTGAACAGTGAAGGATTTAAGAGAAAAATCTTTGTTAATTTGTAGCAACA 3690  
3001 Db TTCAGATGTTTGAACAGTGAAGGATTTAAGAGAAAAATCTTTGTTAATTTGTAGCAACA 3060  
3691 QY CATCAGGATCGGAATCTTATTAAGAAATGATGAGATTTACTCATGATAAAGAAAAAGCAGAAC 3750  
3061 Db CATCAGGATCGGAATCTTATTAAGAAATGATGAGATTTACTCATGATAAAGAAAAAGCAGAAC 3120  
3751 QY GAAAAAGAAAAAGCTGAGCTCTAGGCTACATGCCAGAGATCATGCTCAGATGTCTG 3810  
3121 Db GAAAAAGAAAAAGCTGAGCTCTAGGCTACATGCCAGAGATCATGCTCAGATGTCTG 3180  
3811 QY CCTTACAGAAAACTTCAATTTGAAAATCTATAAACTCATGTTATGACAAATATCATCAGAAATGC 3870  
3181 Db CCTTACAGAAAACTTCAATTTGAAAATCTATAAACTCATGTTATGACAAATATCATCAGAAATGC 3240  
3871 QY CTGGGAAGAAAGATTTCCATTTATGGAGAGAGACCCCGAGAGTCTAGTACTCTCTA 3930  
3241 Db CTGGGAAGAAAGATTTCCATTTATGGAGAGAGACCCCGAGAGTCTAGTACTCTCTA 3300  
3931 QY GAAATGCTTTGGGCTTAAACCGGGTCCATCTGTTTACTGAAAAAGAGGTGCTGACGTGCA 3990  
3301 Db GAAATGCTTTGGGCTTAAACCGGGTCCATCTGTTTACTGAAAAAGAGGTGCTGACGTGCA 3360  
3991 QY TCCTTTGCCAAGAGAACAGAGGTGAAAAATAGAAAAATTAATGCCATGTTATTCGGCT 4050



Db 3361 TCCTTTGCCAAGAGAACAGGAGGTGAAAAATAGAAAAATAATGCCATGGTATTATCGGCCT 3420  
Qy 4051 GTGTCCAGAAATCTACTGCTTAACCCAGCAGCAGGGGAAAAACCCATAGAACTCTCAGGAG 4110  
Db 3421 GTGTCCAGAAATCTACTGCTTAACCCAGCAGCAGGGGAAAAACCCATAGAACTCTCAGGAG 3480  
Qy 4111 AAGCCCTAGAACCCACTTTTCATGGATCCAGACTTGGCATATGGAACTTATACAGGAAGCT 4170  
Db 3481 AAGCCCTAGAACCCACTTTTCATGGATCCAGACTTGGCATATGGAACTTATACAGGAAGCT 3540  
Qy 4171 GTGTGATGTAATGCAACGACAGTGTCTGGCAGAAATATTTTGAAGCTGTACAGCTGAGCT 4230  
Db 3541 GTGGTCATGTAATGCAACGACAGTGTCTGGCAGAAATATTTTGAAGCTGTACAGCTGAGCT 3600  
Qy 4231 CTCAGCAGCGCATTCATGTTGACCTTTTGTGACTTGGAAAGTGGAAATATCTTTGCCCTC 4290  
Db 3601 CTCAGCAGCGCATTCATGTTGACCTTTTGTGACTTGGAAAGTGGAAATATCTTTGCCCTC 3660  
Qy 4291 TTTGCAAAATCTCTGTGCAATACGTGTATCCCATTTATCTTTTGCAACTCAAAAGATAA 4350  
Db 3661 TTTGCAAAATCTCTGTGCAATACGTGTATCCCATTTATCTTTTGCAACTCAAAAGATAA 3720  
Qy 4351 ACAGTGAGATGACAGATCTCTTGTCTCAACTTTTGAACCTGGCAGCGTGGATACAGACTG 4410  
Db 3721 ACAGTGAGATGACAGATCTCTTGTCTCAACTTTTGAACCTGGCAGCGTGGATACAGACTG 3780  
Qy 4411 TTCTGGCCAGAAATATCAGGTTATATATATAGACATGCTAAAGGAGAAACCCAAATCTCTA 4470  
Db 3781 TTCTGGCCAGAAATATCAGGTTATATATATAGACATGCTAAAGGAGAAACCCAAATCTCTA 3840  
Qy 4471 TTTTCTTTTAAACAGGAATGGAGATCTCTTTGGAGTTCATTTCCATCTCGAGTTTG 4530  
Db 3841 TTTTCTTTTAAACAGGAATGGAGATCTCTTTGGAGTTCATTTCCATCTCGAGTTTG 3900  
Qy 4531 GCGTTGAGTCTTCGATTAATAATCAAAATAGCATCAAGGAAATGGTTATTTCTTTTGCCA 4590  
Db 3901 GCGTTGAGTCTTCGATTAATAATCAAAATAGCATCAAGGAAATGGTTATTTCTTTTGCCA 3960  
Qy 4591 CAACAAATTAATAGAAATGGATTTGAAGTGCACCTGATGAAAGGATCTCGAGTCCCA 4650  
Db 3961 CAACAAATTAATAGAAATGGATTTGAAGTGCACCTGATGAAAGGATCTCGAGTCCCA 4020  
Qy 4651 TGCTGACCTGGAGCACTCGCGTTTCACTATCCAGGCAATGAAATCTTATTTGGAGATG 4710  
Db 4021 TGCTGACCTGGAGCACTCGCGTTTCACTATCCAGGCAATGAAATCTTATTTGGAGATG 4080  
Qy 4711 AAGGAAAACTCTGTTTGGAGCACTTCAAAATAGGCAGCATAAATGGTCTGAAAGCATTA 4770  
Db 4081 AAGGAAAACTCTGTTTGGAGCACTTCAAAATAGGCAGCATAAATGGTCTGAAAGCATTA 4140  
Qy 4771 TGCAGTTTGCAGTTGCAACAGAGATTAATCTGTCTCAGGTCCTGATACAGAAACATCTGG 4830  
Db 4141 TGCAGTTTGCAGTTGCAACAGAGATTAATCTGTCTCAGGTCCTGATACAGAAACATCTGG 4200  
Qy 4831 TTGCTCTCTCATCAGTGTCTTCTCAATACATAAATCAGAGATACACCATGCCCTTCGT 4890  
Db 4201 TTGCTCTCTCATCAGTGTCTTCTCAATACATAAATCAGAGATACACCATGCCCTTCGT 4260  
Qy 4891 CTATAGATCTGTTTCATGTTTGGTGGTGTGTGTAGCATTCCTCATCTTTGATTTGGG 4950  
Db 4261 CTATAGATCTGTTTCATGTTTGGTGGTGTGTGTAGCATTCCTCATCTTTGATTTGGG 4320  
Qy 4951 ATGACCTTGTGATCTGACGCTTCTTCAGTTAGTTCCTCAATACCACTTTATCTCT 5010  
Db 4321 ATGACCTTGTGATCTGACGCTTCTTCAGTTAGTTCCTCAATACCACTTTATCTCT 4380  
Qy 5011 TCCATTTGATCACCATGACACATGCTTCAGATACTACTTACAGTAGACACAGCCCTAC 5070  
Db 4381 TCCATTTGATCACCATGACACATGCTTCAGATACTACTTACAGTAGACACAGCCCTAC 4440  
Qy 5071 CCCTTGCTCAGGTTCAAGAGACAGTGAAGAGGCTCATTTCCGCTCTTTCTTTTGCGAG 5130  
Db 4441 CCCTTGCTCAGGTTCAAGAGACAGTGAAGAGGCTCATTTCCGCTCTTTCTTTTGCGAG 4500

Qy 5131 AAATTTCTCAATATACAAGTGGCTCCATTTGGTGTGATATTTCTGGCTGGTATTTTTGGG 5190  
Db 4501 AAATTTCTCAATATACAAGTGGCTCCATTTGGTGTGATATTTCTGGCTGGTATTTTTGGG 4560  
Qy 5191 TCTCACTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGTGCTATTTGTTTCCACTATT 5250  
Db 4561 TCTCACTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGTGCTATTTGTTTCCACTATT 4620  
Qy 5251 TACTTGGGGTAACTCCCGCTGAGGAACATGCAATATCTGCAAGAGGAGATACAGTG 5310  
Db 4621 TACTTGGGGTAACTCCCGCTGAGGAACATGCAATATCTGCAAGAGGAGATACAGTG 4680  
Qy 5311 CACTCTGTAGCTATCTATCTTTTACTCAAAATTTGTTCTGCTCTTCCAGGAATATTGGG 5370  
Db 4681 CACTCTGTAGCTATCTATCTTTTACTCAAAATTTGTTCTGCTCTTCCAGGAATATTGGG 4740  
Qy 5371 ATACTGTAAGGCCCTTGTCTCCAGAGTGGTGTGAGATCTGCGCTTACTAACTGTTTGA 5430  
Db 4741 ATACTGTAAGGCCCTTGTCTCCAGAGTGGTGTGAGATCTGCGCTTACTAACTGTTTGA 4800  
Qy 5431 AGCAAAAAACACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATGAGACTTCTG 5490  
Db 4801 AGCAAAAAACACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATGAGACTTCTG 4860  
Qy 5491 ATGACTATAGCTGCTCTCTGAATCAAGCTTCTCAATTTCAAGTGCACGCTCTGCAGATG 5550  
Db 4861 ATGACTATAGCTGCTCTCTGAATCAAGCTTCTCAATTTCAAGTGCACGCTCTGCAGATG 4920  
Qy 5551 ATGAGCGAAAGCATCTCTGCTGCTCTTTTCTGTGGGCTTATCTATGTTTCTCAGAAC 5610  
Db 4921 ATGAGCGAAAGCATCTCTGCTGCTCTTTTCTGTGGGCTTATCTATGTTTCTCAGAAC 4980  
Qy 5611 TTTGCTGCGAGAAATTTGTGAACGCGGGAAGAGTTGAGCTTGTGATTTTTCACGCACTTC 5670  
Db 4981 TTTGCTGCGAGAAATTTGTGAACGCGGGAAGAGTTGAGCTTGTGATTTTTCACGCACTTC 5040  
Qy 5671 ACTGTGAGCGGAGTCTGCAATTTTCTTAAATATCAGAGATGCGAGTGTCTTGGTTG 5730  
Db 5041 ACTGTGAGCGGAGTCTGCAATTTTCTTAAATATCAGAGATGCGAGTGTCTTGGTTG 5100  
Qy 5731 AAGTAAAGCAGAGGCTGTGCTTATCCAGCTCTTACTTGGATGAATATGAGAGAACAG 5790  
Db 5101 AAGTAAAGCAGAGGCTGTGCTTATCCAGCTCTTACTTGGATGAATATGAGAGAACAG 5160  
Qy 5791 ACCCTGGCTGAGAGGGGCAACCCCTTCAATTTATCTGTGAGGCTATCGAGCTCC 5850  
Db 5161 ACCCTGGCTGAGAGGGGCAACCCCTTCAATTTATCTGTGAGGCTATCGAGCTCC 5220  
Qy 5851 ATTTGCTGTGCAACACATGCAATATAGAGAGATTTGTAGGAGCCCAAGAGACTAATC 5910  
Db 5221 ATTTGCTGTGCAACACATGCAATATAGAGAGATTTGTAGGAGCCCAAGAGACTAATC 5280  
Qy 5911 AGATGTTATTTGGATTTCAACTGGCAGTTTACTGTGAGCTCCAACTCTGCTCAAGCAATC 5970  
Db 5281 AGATGTTATTTGGATTTCAACTGGCAGTTTACTGTGAGCTCCAACTCTGCTCAAGCAATC 5340  
Qy 5971 ACAAATGACGACAGTAGTAAGGCTGATTTCAAAATTTATGAAAACTTTCTGAGGGCTGG 6030  
Db 5341 ACAAATGACGACAGTAGTAAGGCTGATTTCAAAATTTATGAAAACTTTCTGAGGGCTGG 5400  
Qy 6031 AAGCTATTTGAGGCTCTTTTGTCTCCATGCTCAGGTTCACTTACATCAATAAATATTTCT 6090  
Db 5401 AAGCTATTTGAGGCTCTTTTGTCTCCATGCTCAGGTTCACTTACATCAATAAATATTTCT 5460  
Qy 6091 TAATGG 6096  
Db 5461 TAATGG 5466

```
; Publication No. US20040185037A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966B
; CURRENT APPLICATION NUMBER: US/10/758,672A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 5205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (662)..(662)
; OTHER INFORMATION: n = a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (668)..(668)
; OTHER INFORMATION: n = a, c, g, or t
; US-10-758-672A-18

Query Match      81.4%; Score 5136; DB 19; Length 5205;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;

QY 696 ATGGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGGTACCCAG 755
DB 1 ATGGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGGTACCCAG 60
QY 756 ACCCTCAGCGTCTGGCAATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 815
DB 61 ACCCTCAGCGTCTGGCAATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 120
QY 816 CATCATTTGGCAATTTGGTGCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 875
DB 121 CATCATTTGGCAATTTGGTGCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 180
QY 876 AAGCAGGAGGAAAGTGTAACAATGTCAATATTCATCCACTGGAATGTACTTATTTGGA 935
DB 181 AAGCAGGAGGAAAGTGTAACAATGTCAATATTCATCCACTGGAATGTACTTATTTGGA 240
QY 936 GAAGATCCAGATATTTGCTTAGAGAAATTTAGACACAGTGGAGCATTTTCAGCTTTGGG 995
DB 241 GAAGATCCAGATATTTGCTTAGAGAAATTTAGACACAGTGGAGCATTTTCAGCTTTGGG 300
QY 996 AGGGTTTTCAAAGTGGAGAGACAACTTATTTCTGAGGATTTGTGCAATTTGATCCAACA 1055
DB 301 AGGGTTTTCAAAGTGGAGAGACAACTTATTTCTGAGGATTTGTGCAATTTGATCCAACA 360
QY 1056 TGTGTACTCTGTATGAGACTGCTTCAGGACAGTGTTCATAAAATTCATCGTTACAAGATG 1115
DB 361 TGTGTACTCTGTATGAGACTGCTTCAGGACAGTGTTCATAAAATTCATCGTTACAAGATG 420
QY 1116 CATACTTCTACTGGAGGGGTTCTGTGACTGTGTGAGACACAGAGGCATGGAATCTGGC 1175
DB 421 CATACTTCTACTGGAGGGGTTCTGTGACTGTGTGAGACACAGAGGCATGGAATCTGGC 480
QY 1176 CCTTTTGTGTAAATCATGAACCTGGAAGAGCAGGTACTTATAAAGAGAAATTCAGCTGT 1235
DB 481 CCTTTTGTGTAAATCATGAACCTGGAAGAGCAGGTACTTATAAAGAGAAATTCAGCTGT 540
QY 1236 CGTTGGAATGAAGAGGTAAATTTGTCAGCCAGGAAATATTTTCTTTCAGTGTATAAATAT 1295
DB 541 CGTTGGAATGAAGAGGTAAATTTGTCAGCCAGGAAATATTTTCTTTCAGTGTATAAATAT 600
QY 1296 GTCGTAGAAATGACTATATGGGAAGAGAAAGAACTGGCTCTCTGAACTCCAGATAGG 1355
; 601 GTCGTAGAAATGACTATATGGGAAGAGAAAGAAAGAACTGGCTCTCTGAACTCCAGATAAGG 660
; 1356 GAGAAAAATGAAGATACTATTTGTCTCTTTTCAATGATGAACACCAATTCATATGACCAC 1415
; 661 KNR-----YVCVNDHSHVDH 675
; 1416 GTCATATACAGCCTACAAAGAGCTTTGACTGTGTGAGCTCGCAGAGGCGCCAGTTGCATACC 1475
; 676 GTCATATACAGCCTACAAAGAGCTTTGACTGTGTGAGCTCGCAGAGGCGCCAGTTGCATACC 735
; 1476 ACTGCCATTGACAAAGAGGCTGTCGGGCTGTTAAAGCGGGAGCTTATGCTCTCTGCGCAG 1535
; 736 ACTGCCATTGACAAAGAGGCTGTCGGGCTGTTAAAGCGGGAGCTTATGCTCTCTGCGCAG 795
; 1536 GAAGCAAGGAAGATATTAAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 1595
; 796 GAAGCAAGGAAGATATTAAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 855
; 1596 GAAGTATTACACTCAGAGATTATGCTCATCAGAAAAATTTGCTTTGCGCTCTGGTTCTCTGG 1655
; 856 GAAGTATTACACTCAGAGATTATGCTCATCAGAAAAATTTGCTTTGCGCTCTGGTTCTCTGG 915
; 1656 ATGAACAAAATTATGAGCTATTTCAAGTGACTTTAGGCGAGATCTTTTGCCAGCATGCCCTT 1715
; 916 ATGAACAAAATTATGAGCTATTTCAAGTGACTTTAGGCGAGATCTTTTGCCAGCATGCCCTT 975
; 1716 AGAGAAGAACCTGACTCGGAGAAATCCCTGTCTCTCAAGCAGGTAAATGCTTTGGGATGCA 1775
; 976 AGAGAAGAACCTGACTCGGAGAAATCCCTGTCTCTCAAGCAGGTAAATGCTTTGGGATGCA 1035
; 1776 AAGCTTTATAAGGTGCCGTAAGATCCTTCATGATTTGATCTTCAGCAGATTTTTTTATG 1835
; 1036 AAGCTTTATAAGGTGCCGTAAGATCCTTCATGATTTGATCTTCAGCAGATTTTTTTATG 1095
; 1836 GAGATGGAATACAAAAAATCTTTGCTATGAAATTTGTGAAGATATTAATAACAACCTGAG 1895
; 1096 GAGATGGAATACAAAAAATCTTTGCTATGAAATTTGTGAAGATATTAATAACAACCTGAG 1155
; 1896 AAAGAATATATCAGTGATGATCATGACAGAAAGTATCTCTATAACTGCACTTTTCAGTTGAG 1955
; 1156 AAAGAATATATCAGTGATGATCATGACAGAAAGTATCTCTATAACTGCACTTTTCAGTTGAG 1215
; 1956 ATGTTTACTGTTCTTACTCTGCTCGACATCTTATTTGAAGACAGAAATGTTATCTCTGTC 2015
; 1216 ATGTTTACTGTTCTTACTCTGCTCGACATCTTATTTGAAGACAGAAATGTTATCTCTGTC 1275
; 2016 ATTACTGAAACTCTGCTAGAGTTTACCTCAGTACTTTGGACAGAAACAATAAATTTCAAC 2075
; 1276 ATTACTGAAACTCTGCTAGAGTTTACCTCAGTACTTTGGACAGAAACAATAAATTTCAAC 1335
; 2076 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTAAATATGTGACCTAAG 2135
; 1336 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTAAATATGTGACCTAAG 1395
; 2136 TATATCTCTGATCAGCAAAACCCCAATATGGAAGAGATTAAGAATGCAAGTTCTCTGAA 2195
; 1396 TATATCTCTGATCAGCAAAACCCCAATATGGAAGAGATTAAGAATGCAAGTTCTCTGAA 1455
; 2196 GGTTCCTGATCTTTTGAAGATTTCTTACCTGATGCGGGAATGGAAGAAATCCGAAGA 2255
; 1456 GGTTCCTGATCTTTTGAAGATTTCTTACCTGATGCGGGAATGGAAGAAATCCGAAGA 1515
; 2256 CAGGTGGGCAACACATTTGAAGTGGATCTGATTTGGGAGGCTGCGCATTTGCTATACAGATG 2315
; 1516 CAGGTGGGCAACACATTTGAAGTGGATCTGATTTGGGAGGCTGCGCATTTGCTATACAGATG 1575
; 2316 CAATTTGAAGATAATTTTACTCATCTGTCAGAGTGGTGTCTGTGTATGGAAGAACTCTTTA 2375
; 1576 CAATTTGAAGATAATTTTACTCATCTGTCAGAGTGGTGTCTGTGTATGGAAGAACTCTTTA 1635
; 2376 CTCTGCTTATTAAGAAATGTCAAAAGCTGTGTGATGAGTGCAGTACCAAGTTTCTATCT 2435
; 1636 CTCTGCTTATTAAGAAATGTCAAAAGCTGTGTGATGAGTGCAGTACCAAGTTTCTATCT 1695
```

QY 2436 AGTAGAAGACAGTAGTAACTCGTGTGGACATAGTTTGGAAAAACAAGTCTCTACAGAGTA 2495  
DB 1696 AGTAGAAGACAGTAGTAACTCGTGTGGACATAGTTTGGAAAAACAAGTCTCTACAGAGTA 1755  
QY 2496 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTCAAGTA 2555  
DB 1756 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTCAAGTA 1815  
QY 2556 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACGTGCATGAATTTGTGTCTTTTGGAGACTTTT 2615  
DB 1816 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACGTGCATGAATTTGTGTCTTTTGGAGACTTTT 1875  
QY 2616 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTGCTCTGCTGTTGGTGGCCAGGTGCTT 2675  
DB 1876 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTGCTCTGCTGTTGGTGGCCAGGTGCTT 1935  
QY 2676 GCTGAGATGTGGCGAAGAAATGGAGCTGTCTCTTATTAGCCAGGTGTTTATTACCAAGAT 2735  
DB 1936 GCTGAGATGTGGCGAAGAAATGGAGCTGTCTCTTATTAGCCAGGTGTTTATTACCAAGAT 1995  
QY 2736 GTTAAGTGCAGAGAAAGTAATGTATGAATAAGATATCATCTGCTTCAGATTTGGTGCATCT 2795  
DB 1996 GTTAAGTGCAGAGAAAGTAATGTATGAATAAGATATCATCTGCTTCAGATTTGGTGCATCT 2055  
QY 2796 TTAAATGGATCCCAATAAGTCTCTGTTTACTGTTACTTTCAGAGGTATGAACCTTGGCCAGGCT 2855  
DB 2056 TTAAATGGATCCCAATAAGTCTCTGTTTACTGTTACTTTCAGAGGTATGAACCTTGGCCAGGCT 2115  
QY 2856 TTAAACAAGACCATATCTACAAAAGCAGAGATTTGATTAAACAATAATAACACTAATA 2915  
DB 2116 TTAAACAAGACCATATCTACAAAAGCAGAGATTTGATTAAACAATAATAACACTAATA 2175  
QY 2916 GAAGAAATGCTTCAGGTCTCATCTATATTTGTGGTGAAGCTTATGTACCTGGAGTGGGA 2975  
DB 2176 GAAGAAATGCTTCAGGTCTCATCTATATTTGTGGTGAAGCTTATGTACCTGGAGTGGGA 2235  
QY 2976 AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTGCTTTGCAATTTGAACCC 3035  
DB 2236 AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTGCTTTGCAATTTGAACCC 2295  
QY 3036 ATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGAAATGAATAATAATGAACCTGGCTTA 3095  
DB 2296 ATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGAAATGAATAATAATGAACCTGGCTTA 2355  
QY 3096 GAGAATGCTATAAA CAAGTGGCCACATTTAAGAAAACAGGTGTATCAGGCCATGGAGTT 3155  
DB 2356 GAGAATGCTATAAA CAAGTGGCCACATTTAAGAAAACAGGTGTATCAGGCCATGGAGTT 2415  
QY 3156 TATGAATCTAAAGATGAATCACTGAAGACTTCAATATGTACTTTTATCATTTACTCCAAA 3215  
DB 2416 TATGAATCTAAAGATGAATCACTGAAGACTTCAATATGTACTTTTATCATTTACTCCAAA 2475  
QY 3216 ACCCAGCATAGCAAGGCTGAACATATGAGAGAAAAGGAGAAAACAAGAAACAAAGAT 3275  
DB 2476 ACCCAGCATAGCAAGGCTGAACATATGAGAGAAAAGGAGAAAACAAGAAACAAAGAT 2535  
QY 3276 GAAGCATTTGCCCCACCACTCTCTGAAATTTCTGCCCTGCTTTTCAAGAAAGTGAATTAAC 3335  
DB 2536 GAAGCATTTGCCCCACCACTCTCTGAAATTTCTGCCCTGCTTTTCAAGAAAGTGAATTAAC 2595  
QY 3336 CTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTTGGGGGCAATAGAC 3395  
DB 2596 CTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTTGGGGGCAATAGAC 2655  
QY 3396 ACAGATTTCACTGTGACCAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTG 3455  
DB 2656 ACAGATTTCACTGTGACCAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTG 2715  
QY 3456 GGTTTACTAGAGAGAGCAACAGCTTTCAAAAAGCTCTCTGAAAGAGAGTAAATTTGAC 3515  
DB 2716 GGTTTACTAGAGAGAGCAACAGCTTTCAAAAAGCTCTCTGAAAGAGAGTAAATTTGAC 2775

QY 3516 TTTTATCATAGGCTTCAAGATTTGGGAAGTTCCGCCATGAATATACAAATGCTTTTGGAA 3575  
DB 2776 TTTTATCATAGGCTTCAAGATTTGGGAAGTTCCGCCATGAATATACAAATGCTTTTGGAA 2835  
QY 3576 AAATCTAAAGGAATTTCCCAAGTTTGAAGGCCAGAGGACATGATAAAGTGTGATCTTTCAG 3635  
DB 2836 AAATCTAAAGGAATTTCCCAAGTTTGAAGGCCAGAGGACATGATAAAGTGTGATCTTTCAG 2895  
QY 3636 ATGTTTTCACACAGTGAAGCGAATTAAGAGAAAATCTTTGTTTAAATTTGTAGCAACCAATCA 3695  
DB 2896 ATGTTTTCACACAGTGAAGCGAATTAAGAGAAAATCTTTGTTTAAATTTGTAGCAACCAATCA 2955  
QY 3696 GGATCGGAATCTTATTAAGAAATGATGAGATTAATCTCATGATTAAGAAAAGCAAGAA 3755  
DB 2956 GGATCGGAATCTTATTAAGAAATGATGAGATTAATCTCATGATTAAGAAAAGCAAGAA 3015  
QY 3756 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAAAGATCATGGCTCAGATGTCTGCCCTTA 3815  
DB 3016 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAAAGATCATGGCTCAGATGTCTGCCCTTA 3075  
QY 3816 CAGAAAATCTTCAATTTGAAACTCATAAACTCATGATGATGACAATACATCAGAAATGCTCGGG 3875  
DB 3076 CAGAAAATCTTCAATTTGAAACTCATAAACTCATGATGATGACAATACATCAGAAATGCTCGGG 3135  
QY 3876 AAAGAAGATTCATTTATGGAGGAAGAGAGCAACCCAGCAGTCAGTGAATCTCTAGAAAT 3935  
DB 3136 AAAGAAGATTCATTTATGGAGGAAGAGAGCAACCCAGCAGTCAGTGAATCTCTAGAAAT 3195  
QY 3936 GCTTTGGGTCTTAAACGGGTCCATCTGTTTACTGAAAAGGAGGTGCTGACGTGATCTT 3995  
DB 3196 GCTTTGGGTCTTAAACGGGTCCATCTGTTTACTGAAAAGGAGGTGCTGACGTGATCTT 3255  
QY 3996 TGCCAAAGAGAAACAGGAGGTGAAAATAGAAAATTAATGCCATGTTATTTATCGGCTGTGTC 4055  
DB 3256 TGCCAAAGAGAAACAGGAGGTGAAAATAGAAAATTAATGCCATGTTATTTATCGGCTGTGTC 3315  
QY 4056 CAGAAAATCTACTGCTTTAAACCCAGCAGCAAGGGGAAAACCCATAGAACTCTCTCAGGAGAGCC 4115  
DB 3316 CAGAAAATCTACTGCTTTAAACCCAGCAGCAAGGGGAAAACCCATAGAACTCTCTCAGGAGAGCC 3375  
QY 4116 CTAGACCCATCTTTTATGATTCAGACTTGGCAATATGGAATCTTATACAGGAAGCTGTGTT 4175  
DB 3376 CTAGACCCATCTTTTATGATTCAGACTTGGCAATATGGAATCTTATACAGGAAGCTGTGTT 3435  
QY 4176 CATGTAATGACGACAGTGTGCTGCAAGATTAATTTTGAAGCTGTACAGCTCAGCTCTCAG 4235  
DB 3436 CATGTAATGACGACAGTGTGCTGCAAGATTAATTTTGAAGCTGTACAGCTCAGCTCTCAG 3495  
QY 4236 CAGCGCATTCATGTTGACCTTTTGTGACCTTGGAAAAGTGGAGAAATCTTTTGGCTCTTTTC 4295  
DB 3496 CAGCGCATTCATGTTGACCTTTTGTGACCTTGGAAAAGTGGAGAAATCTTTTGGCTCTTTTC 3555  
QY 4296 AAATCTCTGTGCAATATCTGTGATCCCATTTATTTCTTTGCAACCTCAAAAAGATAACAGT 4355  
DB 3556 AAATCTCTGTGCAATATCTGTGATCCCATTTATTTCTTTGCAACCTCAAAAAGATAACAGT 3615  
QY 4356 GAGATGACAGATGCTCTTGTCTCAACTTTTGTGACCTGGCAGCTGGATACAGACTGTTCTG 4415  
DB 3616 GAGATGACAGATGCTCTTGTCTCAACTTTTGTGACCTGGCAGCTGGATACAGACTGTTCTG 3675  
QY 4416 GCCAGAAATATCAGGTTTATAATATTAAGACATGCTTAAAGGAGAAAACCCCAATTTCTTATTTTC 4475  
DB 3676 GCCAGAAATATCAGGTTTATAATATTAAGACATGCTTAAAGGAGAAAACCCCAATTTCTTATTTTC 3735  
QY 4476 TTTAATCAAGAAATGGGAGATTTCTATTTGAGATTTCCATTCCTGAGTTTGGCGCTT 4535  
DB 3736 TTTAATCAAGAAATGGGAGATTTCTATTTGAGATTTCCATTCCTGAGTTTGGCGCTT 3795  
QY 4536 GAGTCTTCGATTAATAATTTCAAAATAGCATCAAGGAAATGGTTATTTCTTTGCCACAACA 4595  
DB 3796 GAGTCTTCGATTAATAATTTCAAAATAGCATCAAGGAAATGGTTATTTCTTTGCCACAACA 3855  
QY 4596 ATTTATAGAAATTTGAAATTTGAAAGTGCACCTGATGAAAGGGATTCCTCGAGTCCCCATGCTG 4655

Db 3856 ATTATAGAAATTTGGATTTGAAAGTGCACCTGATGAAAGGATCCCTCGAGTCCCCATGCTG 3915  
Qy 4656 ACCTGGAGCACCTGGCGCTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATCAAGGA 4715  
Db 3916 ACCTGGAGCACCTGGCGCTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAGGA 3975  
Qy 4716 AAACCTCTGTTTGGAGCACTTCAAAATAGCGAGCAATAATGGTGTGAAAGCATTAATGCG 4775  
Db 3976 AAACCTCTGTTTGGAGCACTTCAAAATAGCGAGCAATAATGGTGTGAAAGCATTAATGCG 4035  
Qy 4776 TTTGAGTGTGACAGAGGATTAACCTGCTCAGGTCTGATACAGAAACATCTGTTGCT 4835  
Db 4036 TTTGAGTGTGACAGAGGATTAACCTGCTCAGGTCTGATACAGAAACATCTGTTGCT 4095  
Qy 4836 CTTCATCAGTTGTTCTTCTTAACATAAAATCAGAGATACACCATGCTCTGCTCTATA 4895  
Db 4096 CTTCATCAGTTGTTCTTCTTAACATAAAATCAGAGATACACCATGCTCTGCTCTATA 4155  
Qy 4896 GATCTGTTTCATGTTTGGTGGTGTGTGATGATCCCATCTTGTATTGGATGAC 4955  
Db 4156 GATCTGTTTCATGTTTGGTGGTGTGTGATGATCCCATCTTGTATTGGATGAC 4215  
Qy 4956 CCGTGTGATCTGACGCTTCTTCACTAGTCTTCTTCTTATTAACCACTTATCTCTCCAT 5015  
Db 4216 CCGTGTGATCTGACGCTTCTTCACTAGTCTTCTTCTTATTAACCACTTATCTCTCCAT 4275  
Qy 5016 TTGATCACCATGGCACACATGCTTCAGATACCTACTTACAGTAGACACAGGCTTACCCCTT 5075  
Db 4276 TTGATCACCATGGCACACATGCTTCAGATACCTACTTACAGTAGACACAGGCTTACCCCTT 4335  
Qy 5076 GCTCAGGTTCAAGAGACAGTGAAGGCTCAATCCGCACTTCTTCTTTTGCAGAAAT 5135  
Db 4336 GCTCAGGTTCAAGAGACAGTGAAGGCTCAATCCGCACTTCTTCTTTTGCAGAAAT 4395  
Qy 5136 TCTCAATATACAGTGGCTCCATTTGGTGTGATATCTCGGCTGTATTTGGTGGTCTCA 5195  
Db 4396 TCTCAATATACAGTGGCTCCATTTGGTGTGATATCTCGGCTGTATTTGGTGGTCTCA 4455  
Qy 5196 CTGAAGATGGCATCACCCCTTATCTTCTGCTGTGCTGATTTCTTCTTCTTCTTCTTCTT 5255  
Db 4456 CTGAAGATGGCATCACCCCTTATCTTCTGCTGTGCTGATTTCTTCTTCTTCTTCTTCTT 4515  
Qy 5256 GGGGTAACCTCGCTGAGGAACTGCATACCAATCTGCAAGAGGAGATACAGTGCACCTC 5315  
Db 4516 GGGGTAACCTCGCTGAGGAACTGCATACCAATCTGCAAGAGGAGATACAGTGCACCTC 4575  
Qy 5316 TGTAGCTATCTATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTGGGATACT 5375  
Db 4576 TGTAGCTATCTATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTGGGATACT 4635  
Qy 5376 GTAAGGCTTGTCTCCAGAGGTGGTGTGCAGATCTGCTTACTTAACTGTTTGAAGCAA 5435  
Db 4636 GTAAGGCTTGTCTCCAGAGGTGGTGTGCAGATCTGCTTACTTAACTGTTTGAAGCAA 4695  
Qy 5436 AAAAACACCGTGGTCAAGTACCTTAGAAGAAAGAAATAGTTGATAGAGCTTCTCATGAC 5495  
Db 4696 AAAAACACCGTGGTCAAGTACCTTAGAAGAAAGAAATAGTTGATAGAGCTTCTCATGAC 4755  
Qy 5496 TATAGCTGCTCTGAAATCAAGCTTCTCATTTTCAAGTGGCCACGCTCTGCGAGATGAG 5555  
Db 4756 TATAGCTGCTCTGAAATCAAGCTTCTCATTTTCAAGTGGCCACGCTCTGCGAGATGAG 4815  
Qy 5556 CGAAGACATCTGCTCTGCTCTTCTGTTGGGCTATCTATGTTCTCAGAACATTTGC 5615  
Db 4816 CGAAGACATCTGCTCTGCTCTTCTGTTGGGCTATCTATGTTCTCAGAACATTTGC 4875  
Qy 5616 TGCCAGGAAATTTGAAACGGGGAAGAGGTTGGAGCTTGCATTTTTCAGCACTTCACTGT 5675  
Db 4876 TGCCAGGAAATTTGAAACGGGGAAGAGGTTGGAGCTTGCATTTTTCAGCACTTCACTGT 4935  
Qy 5676 GGAGCCGAGTCTGCAATTTCTTAAATCAGAGAAATCCGAGTGGTCTCTGGTTGAAGGT 5735

Db 4936 GGAGCCGAGTCTGCAATTTTCTTAAATCAGAGAAATCCGAGTGGTCTCTGGTTGAAGGT 4995  
Qy 5736 AAAGCCAGAGGCTGTGCCCTATCCAGCTCTCTTACTTTGGATGATATCGAGAAACAGACCCCT 5795  
Db 4996 AAAGCCAGAGGCTGTGCCCTATCCAGCTCTCTTACTTTGGATGATATCGAGAAACAGACCCCT 5055  
Qy 5796 GGCCTGAAGAGGGGCAACCCCTTCAATTTATCTCTGAGCGGTATCGGAAGCTCCATTTG 5855  
Db 5056 GGCCTGAAGAGGGGCAACCCCTTCAATTTATCTCTGAGCGGTATCGGAAGCTCCATTTG 5115  
Qy 5856 GTCTGGCAACACACCTGCAATTTATGAAGAGATTTCTAGGAGCAAGAGACTTAATCAGATG 5915  
Db 5116 GTCTGGCAACACACCTGCAATTTATGAAGAGATTTCTAGGAGCAAGAGACTTAATCAGATG 5175  
Qy 5916 TTATTTGATTTCAACTGGCAGTTACTGTGA 5945  
Db 5176 TTATTTGATTTCAACTGGCAGTTACTGTGA 5205

## RESULT 7

US-10-758-636A-18  
; Sequence 18, Application US/10758636A  
; Publication No. US20050089876A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966C  
; CURRENT APPLICATION NUMBER: US/10758,636A  
; PRIORITY FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 5205  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (662)..(662)  
; OTHER INFORMATION: n = a, c, g, or t  
; NAME/KEY: misc feature  
; LOCATION: (668)..(668)  
; OTHER INFORMATION: n = a, c, g, or t  
US-10-758-636A-18

Query Match 81.4%; Score 5136; DB 21; Length 5205;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;

Qy 696 ATGGCGGACGAGGAGGCTGAGGTAATCTGAGAGGATGGAATCAGCGCGAGGTTACCCAG 755  
Db 1 ATGGCGGACGAGGAGGCTGAGGTAATCTGAGAGGATGGAATCAGCGCGAGGTTACCCAG 60  
Qy 756 ACCCTCAGCTCTGGCAATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 815  
Db 61 ACCCTCAGCTCTGGCAATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 120  
Qy 816 CATCATTTGGCAAAATTTGGTGGGATGGAATTTTCTGCTGAAATGGAATCCAGACTTGGAA 875  
Db 121 CATCATTTGGCAAAATTTGGTGGGATGGAATTTTCTGCTGAAATGGAATCCAGACTTGGAA 180  
Qy 876 AAGCAGGAGGAAAGTGTACAAATGTCAATATTCATCCATCGGAATGTGATTTTGGG 935  
Db 181 AAGCAGGAGGAAAGTGTACAAATGTCAATATTCATCCATCGGAATGTGATTTTGGG 240  
Qy 936 GAAGATCCAGATATTTGCTTAGAGAAATTTGAAGCAGAGTGGAGCAATTTTCACTTTGCGG 995  
Db 241 GAAGATCCAGATATTTGCTTAGAGAAATTTGAAGCAGAGTGGAGCAATTTTCACTTTGCGG 300





Db 2416 TATGAACATAAAGATGAATCACTCAAGAGCTTCAATATATGTACTTTTATCATTTACTTCCAAA 2475  
QY 3216 ACCGAGCATPAGCAGGCTGAACATATGCGAAGAAAGAGAGAAAACAAGAAAACAAGAT 3275  
Db 2476 ACCGAGCATPAGCAGGCTGAACATATGCGAAGAAAGAGAGAAAACAAGAAAACAAGAT 2535  
QY 3276 GAAGCATTTGCCGCCACCAACCACTCTGAAATTTCTGCCCTGCTTTTTCAGCAGAGTGAATTAAC 3335  
Db 2536 GAAGCATTTGCCGCCACCAACCACTCTGAAATTTCTGCCCTGCTTTTTCAGCAGAGTGAATTAAC 2595  
QY 3336 CTTCCTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTTGGCGGGCAATAGAC 3395  
Db 2596 CTTCCTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTTGGCGGGCAATAGAC 2655  
QY 3396 ACAGATTTCTAATCTGTGACCGAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGCAATG 3455  
Db 2656 ACAGATTTCTAATCTGTGACCGAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGCAATG 2715  
QY 3456 GGTTTACTAGAGAGAGAGCAACAGCTTCAAAAAGCTTCTGAAAGAGTCTTGAAGAGAGTAACTTTGAC 3515  
Db 2716 GGTTTACTAGAGAGAGAGCAACAGCTTCAAAAAGCTTCTGAAAGAGTCTTGAAGAGAGTAACTTTGAC 2775  
QY 3516 TTTTATCATTAAGGCTTCAAGATTGGGAAGTTCAAGCATGAATATATCAAAATGCTTTTGGAA 3575  
Db 2776 TTTTATCATTAAGGCTTCAAGATTGGGAAGTTCAAGCATGAATATATCAAAATGCTTTTGGAA 2835  
QY 3576 AAATCAAGAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATTAACGTGGATCTTCAG 3635  
Db 2836 AAATCAAGAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATTAACGTGGATCTTCAG 2895  
QY 3636 ATGTTTGACACAGTCAAGGGATTAGAGAAAATCTTGTTTAAATTTGTAGCAACCAATCA 3695  
Db 2896 ATGTTTGACACAGTCAAGGGATTAGAGAAAATCTTGTTTAAATTTGTAGCAACCAATCA 2955  
QY 3696 GGATCGGAATCTATTAAAGATGATCAGATTACTCATGATTAAGAAAAGAGAGAGAGAGAAA 3755  
Db 2956 GGATCGGAATCTATTAAAGATGATCAGATTACTCATGATTAAGAAAAGAGAGAGAGAGAAA 3015  
QY 3756 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3815  
Db 3016 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3075  
QY 3816 CAGAAAACCTTCAATGAACTCATTAACCTCATGATGACATATACATACAGAAATGCTCGGG 3875  
Db 3076 CAGAAAACCTTCAATGAACTCATTAACCTCATGATGACATATACATACAGAAATGCTCGGG 3135  
QY 3876 AAAGAAGATTTCCATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3935  
Db 3136 AAAGAAGATTTCCATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3195  
QY 3936 GCTTTGGGCTCTAAACGGGGTCCATCTGTATTGAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 3995  
Db 3196 GCTTTGGGCTCTAAACGGGGTCCATCTGTATTGAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 3255  
QY 3996 TGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4055  
Db 3256 TGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3315  
QY 4056 CAGAAATCTACTGCTTTAAACCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4115  
Db 3316 CAGAAATCTACTGCTTTAAACCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3375  
QY 4116 CTAGACCCATCTTTTCAATGGATCCAGATCTGGCATATGGAATTTATACAGAGAGAGAGAGAG 4175  
Db 3376 CTAGACCCATCTTTTCAATGGATCCAGATCTGGCATATGGAATTTATACAGAGAGAGAGAGAG 3435  
QY 4176 CATGTAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4235  
Db 3436 CATGTAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3495  
QY 4236 CAGGCAATCTCATGTTGACCTTTTGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4295  
Db 3496 CAGGCAATCTCATGTTGACCTTTTGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3555

QY 4296 AAATCTCTGTGCAATFACGTGTGATCCCAATTAATTCCTTTTGCAACCTCAAAAGATAAACAGT 4355  
Db 3556 AAATCTCTGTGCAATFACGTGTGATCCCAATTAATTCCTTTTGCAACCTCAAAAGATAAACAGT 3615  
QY 4356 GAGAAATGAGATGCTCTGCTCACTTTTGACCCCTGCGACGCGTGATACAGACTGCTTCTG 4415  
Db 3616 GAGAAATGAGATGCTCTGCTCACTTTTGACCCCTGCGACGCGTGATACAGACTGCTTCTG 3675  
QY 4416 GCCAGAAATCAGGTTTATAATATAAGACATGCTAAAGAGAGAAAACCCAAATTCCTATTTTTC 4475  
Db 3676 GCCAGAAATCAGGTTTATAATATAAGACATGCTAAAGAGAGAAAACCCAAATTCCTATTTTTC 3735  
QY 4476 TTTAATCAAGAAATGAGAGATTTCTACTTTTGAGATTCATTTCAATCCTCAGATTTTGGCGTT 4535  
Db 3736 TTTAATCAAGAAATGAGAGATTTCTACTTTTGAGATTCATTTCAATCCTCAGATTTTGGCGTT 3795  
QY 4536 GAGTCTTCGATTAATAATTTCAAAATAGCATCAAGGAAATGGTTATTTCTCTTTTGCCCAACAA 4595  
Db 3796 GAGTCTTCGATTAATAATTTCAAAATAGCATCAAGGAAATGGTTATTTCTCTTTTGCCCAACAA 3855  
QY 4596 ATTTATAGAAATGGAATGAAAGTCCACTCATGAAAGGGATCCTCGAGTCCCAATGCTG 4655  
Db 3856 ATTTATAGAAATGGAATGGAAGTCCACTCATGAAAGGGATCCTCGAGTCCCAATGCTG 3915  
QY 4656 ACCTGGAGACCTGCGCTTTCATATCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGA 4715  
Db 3916 ACCTGGAGACCTGCGCTTTCATATCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGA 3975  
QY 4716 AAACCTCTGTTGGAGACATTTCAAAATAGGAGATTAATGCTCTGAAAGCATTAATGAG 4775  
Db 3976 AAACCTCTGTTGGAGACATTTCAAAATAGGAGATTAATGCTCTGAAAGCATTAATGAG 4035  
QY 4776 TTTGAGTTGACAGAGATTAACCTCTCAGGCTCTGATACAGAAACATCTGCTTCTG 4835  
Db 4036 TTTGAGTTGACAGAGATTAACCTCTCAGGCTCTGATACAGAAACATCTGCTTCTG 4095  
QY 4836 CTTCATCAGTTGTTCTTCTTAAACATATAAATCAGAAATACACCAATGCTTCTGCTATA 4895  
Db 4096 CTTCATCAGTTGTTCTTCTTAAACATATAAATCAGAAATACACCAATGCTTCTGCTATA 4155  
QY 4896 GATCTGTTTCAATGTTTGGTGGGCTGTGTAGATTCCTCATCTGTATTTGGATGAC 4955  
Db 4156 GATCTGTTTCAATGTTTGGTGGGCTGTGTAGATTCCTCATCTGTATTTGGATGAC 4215  
QY 4956 CCTGTTGATCTGACAGCTTCTTCAAGTTAGTTCTTCTTAAACCACTTATCTCTTCCAT 5015  
Db 4216 CCTGTTGATCTGACAGCTTCTTCAAGTTAGTTCTTCTTAAACCACTTATCTCTTCCAT 4275  
QY 5016 TTGATCAACATGGCACATGCTTTCAGATACTTTCAGTAGACACAGGCTTACCCCTT 5075  
Db 4276 TTGATCAACATGGCACATGCTTTCAGATACTTTCAGTAGACACAGGCTTACCCCTT 4335  
QY 5076 GCTCAGGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5135  
Db 4336 GCTCAGGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4395  
QY 5136 TCTCAATATACAGTGGCTCCATTTGGTGGTGTGATATCTCTGCTGTGATTTTGTGGTGTCTCA 5195  
Db 4396 TCTCAATATACAGTGGCTCCATTTGGTGGTGTGATATCTCTGCTGTGATTTTGTGGTGTCTCA 4455  
QY 5196 CTGAAGAAATGGCATCACCCCTTATCTTGGCTGTGCTGATTTTGTGGTGTGATTTTGTGGTGTCT 5255  
Db 4456 CTGAAGAAATGGCATCACCCCTTATCTTGGCTGTGCTGATTTTGTGGTGTGATTTTGTGGTGTCT 4515  
QY 5256 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 5315  
Db 4516 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 4575  
QY 5316 TGTAGCTATCTATCTTTACCTACAAATTTGCTTCTGCTCTTCCAGGATATTTGGGATACT 5375  
Db 4576 TGTAGCTATCTATCTTTACCTACAAATTTGCTTCTGCTCTTCCAGGATATTTGGGATACT 4635



Db 901 ACATCCACTTCATGTAGAGTATTACACTCAGAGATTATGGCTCATCAGAAATTTGCTTT 960  
Qy 1640 GCGTCTTGGTTCCTGGATGAACAAATTAATAGACTATTCAAGTACTTTAGGACAGATCTT 1699  
Db 961 GCGTCTTGGTTCCTGGATGAACAAATTAATAGACTATTCAAGTACTTTAGGACAGATCTT 1020  
Qy 1700 TTGCCAAGCATGCCCTTAGGAGAACCTTGACTCGAGAAATCCCTGCTCATAAGCAGTT 1759  
Db 1021 TTGCCAAGCATGCCCTTAGGAGAACCTTGACTCGAGAAATCCCTGCTCATAAGCAGTT 1080  
Qy 1760 AATGCTTTGGGATGCAAAAGCTTTATAAAGGTGCGGTAAAGTCTTTCAATGAATTCATCTT 1819  
Db 1081 AATGCTTTGGGATGCAAAAGCTTTATAAAGGTGCGGTAAAGTCTTTCAATGAATTCATCTT 1140  
Qy 1820 CAGCAGTTTTTTTATGGAGATGGAATACAAAACTCTTTGCTATGGAATTTGGAAGTA 1879  
Db 1141 CAGCAGTTTTTTTATGGAGATGGAATACAAAACTCTTTGCTATGGAATTTGGAAGTA 1200  
Qy 1880 TTATAAACAATGCAAGAAATATATCAGTGTATGATCATGACAGAAATCTCTATTAAC 1939  
Db 1201 TTATAAACAATGCAAGAAATATATCAGTGTATGATCATGACAGAAATCTCTATTAAC 1260  
Qy 1940 TGCACTTTCACTTCAGATGTTTACTGTTCTCTACTCTGCTCGACATCTTATTGAAGACA 1999  
Db 1261 TGCACTTTCACTTCAGATGTTTACTGTTCTCTACTCTGCTCGACATCTTATTGAAGACA 1320  
Qy 2000 GAAATGTTATCTCTGTCATTACTGAACTCTGCTAGAAATTTTACCTGAGTACTTTGACAG 2059  
Db 1321 GAAATGTTATCTCTGTCATTACTGAACTCTGCTAGAAATTTTACCTGAGTACTTTGACAG 1380  
Qy 2060 GAACATTAATCACTTCAGGTTTATAGCCAGGACAAATTTGGGAAGATATATGCACT 2119  
Db 1381 GAACATTAATCACTTCAGGTTTATAGCCAGGACAAATTTGGGAAGATATATGCACT 1440  
Qy 2120 AATATGTCACCTTAAAGTATATCTCTATCAGCAAACTTATGAGTATGCAAGAAATTAAG 2179  
Db 1441 AATATGTCACCTTAAAGTATATCTCTATCAGCAAACTTATGAGTATGCAAGAAATTAAG 1500  
Qy 2180 AATGAGTTCTTGAAGTTTTCGATCTTTTGAAGATTTCTTACCTGTATGCAAGGAAAT 2239  
Db 1501 AATGAGTTCTTGAAGTTTTCGATCTTTTGAAGATTTCTTACCTGTATGCAAGGAAAT 1560  
Qy 2240 GGAAGAAATCCGAGACAGGTTGGGCAACATTTGAAGTGTATCTGATGGAGGCTGC 2299  
Db 1561 GGAAGAAATCCGAGACAGGTTGGGCAACATTTGAAGTGTATCTGATGGAGGCTGC 1620  
Qy 2300 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTTCCAAAGAGTGTGCTTG 2359  
Db 1621 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTTCCAAAGAGTGTGCTTG 1680  
Qy 2360 TGATGAAGAACTTTACTTTGGCTTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGCAG 2419  
Db 1681 TGATGAAGAACTTTACTTTGGCTTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGCAG 1740  
Qy 2420 TACAGTTTCAATCTAGTAGCAGACAGTAGTAGTACAATCGTGTGGAACATAGTTTGGAAAC 2479  
Db 1741 TACAGTTTCAATCTAGTAGCAGACAGTAGTAGTACAATCGTGTGGAACATAGTTTGGAAAC 1800  
Qy 2480 AAAGTCTTACAGATGATCTGAGGATCTTTGTAAGCATATCTGCCACTCTCTAGGACCT 2539  
Db 1801 AAAGTCTTACAGATGATCTGAGGATCTTTGTAAGCATATCTGCCACTCTCTAGGACCT 1860  
Qy 2540 TGCTGGTCTTCAATGTACGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGCAATGAATTTGT 2599  
Db 1861 TGCTGGTCTTCAATGTACGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGCAATGAATTTGT 1920  
Qy 2600 GTCTTTTGAAGCATTTCAAGTAGAGGTACTAGTGAATATCTTTAGCTGTCTGCTGCTT 2659  
Db 1921 GTCTTTTGAAGCATTTCAAGTAGAGGTACTAGTGAATATCTTTAGCTGTCTGCTGCTT 1980  
Qy 2660 GGTGGCCAGGTTGTTGCTGAGATGTGGCGAAGAAATGACATGCTCTCTATTAGCCAGGT 2719  
Db 1981 GGTGGCCAGGTTGTTGCTGAGATGTGGCGAAGAAATGACATGCTCTCTATTAGCCAGGT 2040

Qy 2720 GTTTTATTACCAAGATGTTTAAAGTGCAGAGAAAGAAATGTATGATTAAGATATCATCATGCT 2779  
Db 2041 GTTTTATTACCAAGATGTTTAAAGTGCAGAGAAAGAAATGTATGATTAAGATATCATCATGCT 2100  
Qy 2780 TCAGATTGGTGCATCTTTTAAATGGATCCCAATAAGTTCTTGTGTACTGTGACTTTCAGAGTA 2839  
Db 2101 TCAGATTGGTGCATCTTTTAAATGGATCCCAATAAGTTCTTGTGTACTGTGACTTTCAGAGTA 2160  
Qy 2840 TGAACCTTGGCCGAGGCTTTTAAACAAGACCATATCTACAAAAGACCAAGGATTTGATTTAAACA 2899  
Db 2161 TGAACCTTGGCCGAGGCTTTTAAACAAGACCATATCTACAAAAGACCAAGGATTTGATTTAAACA 2220  
Qy 2900 ATATAATACACTAATAGAAAGAAATGCTTCAGGTCTCTCATCTATATTGTGGGTGAGCGTTA 2959  
Db 2221 ATATAATACACTAATAGAAAGAAATGCTTCAGGTCTCTCATCTATATTGTGGGTGAGCGTTA 2280  
Qy 2960 TGTACTCTGGAGTGGAAATGTGACCAAGAGAGGTCACAATGAGAGAAATCAATTCACCTT 3019  
Db 2281 TGTACTCTGGAGTGGAAATGTGACCAAGAGAGGTCACAATGAGAGAAATCAATTCACCTT 2340  
Qy 3020 GCTTTGCAATTCGAACCCCATGCCACACAGTGCCATTTGCCAAAAATTTTACCTGAGAAATGAAAA 3079  
Db 2341 GCTTTGCAATTCGAACCCCATGCCACACAGTGCCATTTGCCAAAAATTTTACCTGAGAAATGAAAA 2400  
Qy 3080 TAATGAAACTGGCTTTAGAGAAATGTCAATAACAAGTGGCCACATTTAAGAAAAACAGGTGT 3139  
Db 2401 TAATGAAACTGGCTTTAGAGAAATGTCAATAACAAGTGGCCACATTTAAGAAAAACAGGTGT 2460  
Qy 3140 ATCAGGCCATGGAGTTTATGAACTTAAAGATGAAATCACTGAAAGCTTCAATATGTACTT 3199  
Db 2461 ATCAGGCCATGGAGTTTATGAACTTAAAGATGAAATCACTGAAAGCTTCAATATGTACTT 2520  
Qy 3200 TTATCATTTACTTCCAAAAACCCAGCATAGCAAGCTTGAACATATGCAAGAAAAAGAGAAA 3259  
Db 2521 TTATCATTTACTTCCAAAAACCCAGCATAGCAAGCTTGAACATATGCAAGAAAAAGAGAAA 2580  
Qy 3260 ACAAGAAAAACAAGATGAAAGCATTTGCCGCCACCAACCTCTCTGAAATTTCTGCCCTGCTTT 3319  
Db 2581 ACAAGAAAAACAAGATGAAAGCATTTGCCGCCACCAACCTCTCTGAAATTTCTGCCCTGCTTT 2640  
Qy 3320 CAGCAAGTGTAACTTCTCACTGTGATATCATGATGTACTTCTCAGGACCGTATTT 3379  
Db 2641 CAGCAAGTGTAACTTCTCACTGTGATATCATGATGTACTTCTCAGGACCGTATTT 2700  
Qy 3380 TGACGGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGCTTCCAAATGGCTTTT 3439  
Db 2701 TGACGGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGCTTCCAAATGGCTTTT 2760  
Qy 3440 TCATATTTCTGGCATTTGGTTTACTAGAGAGAACCAACAGCTTCAAAAAGCTTCTGGAAGA 3499  
Db 2761 TCATATTTCTGGCATTTGGTTTACTAGAGAGAACCAACAGCTTCAAAAAGCTTCTGGAAGA 2820  
Qy 3500 AGAAGTAACTTTGCACTTTTATCATTAAGCTTCAAGATTTGGAAAGTTTTCAGCCATGAATAT 3559  
Db 2821 AGAAGTAACTTTGCACTTTTATCATTAAGCTTCAAGATTTGGAAAGTTTTCAGCCATGAATAT 2880  
Qy 3560 ACAATCTTTTGGAAAAAATCTCAAGGAAATTTCCCAAGTTAGAACGCCAAGAGGACATGAT 3619  
Db 2881 ACAATCTTTTGGAAAAAATCTCAAGGAAATTTCCCAAGTTAGAACGCCAAGAGGACATGAT 2940  
Qy 3620 AACGTGGATACTTCAGATGTTTGAACAGTGAAGCGGATTAAGAGAAAAAATCTTGTGTTAAT 3679  
Db 2941 AACGTGGATACTTCAGATGTTTGAACAGTGAAGCGGATTAAGAGAAAAAATCTTGTGTTAAT 3000  
Qy 3680 TGTAGAACCAACATCAGATCGGAAATCTTATTAAGAAATGAAGATTAATCATGATAAG 3738  
Db 3001 TGTAGAACCAACATCAGATCGGAAATCTTATTAAGAAATGAAGATTAATCATGATAAG 3059

RESULT 9  
US-10-758-672A-3  
; Sequence 3, Application US/10758672A

Publication No. US20040185037A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966B  
; CURRENT APPLICATION NUMBER: US/10/758,672A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 6300  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (295)..(5559)  
US-10-758-672A-3

Query Match 21.9%; Score 1381.8; DB 19; Length 6300;  
Best Local Similarity 56.5%; Pred. No. 0;  
Matches 2947; Conservative 0; Mismatches 2167; Indels 105; Gaps 16;

QY	780	TGGGATCAGCAAGTTGATTTTATTA	1494	GGTCGTCGGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAGCAAGCAAGAGATATA	1553
DB	376	TGGCTGCAAGCAACTGACCTCACTAGAGAAAGTGTTACCAAGCATTTTAGCCCATATGTACCC	1096	GGGCTAGAGTCTGTTCCGATATGGAGATTTTCAGTATTTGTGAGCAAGCAAAATCAGTAAT	1155
QY	840	GAAATTTACTTTGCTGAAATGACCCAGCAGCTTGGAAAAAGCAGGAGAAAGTGTTACAAATG	1554	AAGAGTCATTACAGAAAAATGCTCTCAACATCCACTTCATGTAGAGATTTACACTCAGAG	1613
DB	436	AAATCTACTGCGAGGGTCCCAACCTTTTCCAGAAAGAGACATGCTGGCAGCAT	1156	GTGAGAAATACCACTAGACAGAC---AAAGCCACTCAAGTTCAAGTTATGCAATTCGTCT	1212
QY	900	TCAATTAATCACTCCACTGGAATGGTACTTAATTTGGAGAAGATCCAGATTAATTTGCTTAGAG	1614	ATTATGGCTCATCAGAAATTTGCTTTGGTCTTGGTCTTCCTGGATGAACAAAAATTTATGAGC	1673
DB	496	GTTTTGTGGGACCAATGATGTTGACCTTTGCTGATGAGATCCCTGCAATTTGGATTTTCCA	1213	ATTGTCGCACATCAGAAATTTTGGTTTGAACCTTTTGTCTTGGCTGGGAAGTATTTATGGA	1272
QY	960	AAATTTGAAGCAC---AGTGGAGCATTTTTCAGCTTTTGGGAGGGTTTTCAAAAGTGGAG	1674	TATTCAGATGACTTTTAGGCAGATCTTTTGCACAGCATGCTTTAGCAAGCAACCTGCACTCG	1733
DB	556	AAACTTGAAGCAAGCAAAACCTTCTCATCTTTGCTGATGAGATCCCTGCAATTTGGATTTCCA	1273	TATTCAGATGGCCTTCGCGGATTTTATGTCAAGTTGGTTTACAGAAAGGCCAGATGGT	1332
QY	1017	ACAACCTATTCTTGCAAGGATTTGCAATTTGATCCAACTGATGTTACTCTGTATGACTGC	1734	GAGAAATCCCTGTCTCATTAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCC	1793
DB	616	CTCATATTTCTTGCAAGACATGTCGATGATGATCCAACTTGTGTTTGTGCAATGAGTGC	1333	GAAAACTCTTCTCTAGTGGACAGACTGATGCTTAGTGAATCCAAATTTATGAAAGGTGCT	1392
QY	1077	TTCCAGGACAGTGTTCATAAAATCATCGTTTCAAGATGTCATCTTCTACTGGAGGG	1794	CGTAAGATCCTTTCATGAATTTGATCTTCAGCAGTCTTTCAGGAGTGGATACAAAAA	1853
DB	676	TTTTTGGGAAAGTATTTACAGAGATCATCGATATAGGATGACAAACATCAGGAGTGGAGT	1393	AGGAGTGATATATCATCATCAGTTGTTTCATGAGCAGTCTGCTTATGGATTTGAAATACAGAAA	1452
QY	1137	TTCTGTGACTTGGAGACACAGAGCATGGAACCTGGCCCTTTTGTGTAATCATGAA	1854	CTCTTTGCTATGGAAATTTGTGAAGTATTATAAAACAACTGCAGAAAAAGATATATCAGTGAT	1913
DB	736	TTCTGTGACTGTGATGATCTGAAAGCCCTGGAAGAGGGTCTTACTGTCAAAAACATGAA	1453	CTATTTGCTGTTGCTTTGCAAAAAATTAACGAGTTGTCAGAGAGATTTTATGGAGAT	1512
QY	1197	CTGGAAGACAGGTGACT---ATAAAAGAGAAATTCACGCTGCTGCTGAAGAGCTA	1914	GATCATGACAGAAAGTATCTCTATAAATGCTTTTCAGTTTCAGATGTTTACTGTTCTACT	1973
DB	796	CTTAACACTCTGAAATTTGAGGAGAGAGAGTCTCTTGTTCATTTATCAGAAATGTG	1513	GATCAGCAGCAGCAGTGTGGTGACTCTCTACTGTCCAGTTCTTTCACGCACTACT	1572
QY	1254	ATTGTCCCAAGCCAGGAAATATTTCTTTCAGTGATAAAATATGCTGTAGAAATGACTATA	1974	CTGGCTCGACATCTTATTTGAAGACAGAAATGTTATCTCTGTCATTACTTGAACCTCTGCTA	2033
DB	856	ATAGCAAGAACTTATAACAATTTTGTCTATTACGTTTTCGTTATGCAAGTAAATTAACC	1573	CTGGCTCGAATGCTCATCAACAGAGAAACCTTGAAGACATTATCATTTAAGACTTTTATG	1632
QY	1314	TGGGAAGAGGAAAGAACTGCTCTGAACTTCCAGATTAAGGAGAAATTAAGATATAC	2034	GAAGTTTTT-----ACCTGAGTACTTGCAGAGGAAACAAATAAATTTCAACTTCCAGGGTTAT	2087
DB	916	TGGGAAAAAGAAAGTGAATTTGCCAGCAGATTTAGAGATGGTATAGAAAGTGACACCTAC	1633	GATCATTTGAGACATCGAGATGCCAGGGCAGATTTTCAGTTTGAACGATACACTGCTTTA	1692
QY	1374	TATTTGTCTCTTTTCAATGATCAACACCATTCATATGACCATGTCATATACAGCTTCAA	2088	AGCCAGGACAAATTTGGAGAAGTATATGCAAGTATATGTCACCTTAAAGTATATCTCTGATC	2147
DB	976	TATTTGATGCTGTTTAAAGAGAGGTTTCAACCTATGAAAGATTTATTTACTCTT	1693	CAAGCTTCAAAATTTAGGAGAGTACAGAGCCTTATTTTAGATCTCAAGTATGTGTTAAT	1752
QY	1434	AGAGCTCTTGACTGTGAGCTCGCAGAGGCCCAAGTTGTCATACCACTGCCAATGCAAGAG	2148	AGCAAAACCAATATGGAAGAAAGATTAAGATGCAAGTTCCCTTGAAGGTTTTCGATCT	2207
DB	1036	AAAGCTGTTAACTGTATACAAAAAGAGCTATTGCTTTTGGTTTTCACACTACAGTAGATCGAGAT	1753	AGCAAAACCACTGAATGGTCAGATGAGCTGAGGAGAGATTTCTAGAAAGGGTTTGAATGCC	1812

QY	1494	GGTCGTCGGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAGCAAGCAAGAGATATA	1553
DB	1096	GGGCTAGAGTCTGTTCCGATATGGAGATTTTCAGTATTTGTGAGCAAGCAAAATCAGTAAT	1155
QY	1554	AAGAGTCATTACAGAAAAATGCTCTCAACATCCACTTCATGTAGAGATTTACACTCAGAG	1613
DB	1156	GTGAGAAATACCACTAGACAGAC---AAAGCCACTCAAGTTCAAGTTATGCAATTCGTCT	1212
QY	1614	ATTATGGCTCATCAGAAATTTGCTTTGGTCTTGGTCTTCCTGGATGAACAAAAATTTATGAGC	1673
DB	1213	ATTGTCGCACATCAGAAATTTTGGTTTGAACCTTTTGTCTTGGCTGGGAAGTATTTATGGA	1272
QY	1674	TATTCAGATGACTTTTAGGCAGATCTTTTGCACAGCATGCTTTAGCAAGCAACCTGCACTCG	1733
DB	1273	TATTCAGATGGCCTTCGCGGATTTTATGTCAAGTTGGTTTACAGAAAGGCCAGATGGT	1332
QY	1734	GAGAAATCCCTGTCTCATTAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCC	1793
DB	1333	GAAAACTCTTCTCTAGTGGACAGACTGATGCTTAGTGAATCCAAATTTATGAAAGGTGCT	1392
QY	1794	CGTAAGATCCTTTCATGAATTTGATCTTCAGCAGTCTTTCAGGAGTGGATACAAAAA	1853
DB	1393	AGGAGTGATATATCATCATCAGTTGTTTCATGAGCAGTCTGCTTATGGATTTGAAATACAGAAA	1452
QY	1854	CTCTTTGCTATGGAAATTTGTGAAGTATTATAAAACAACTGCAGAAAAAGATATATCAGTGAT	1913
DB	1453	CTATTTGCTGTTGCTTTGCAAAAAATTAACGAGTTGTCAGAGAGATTTTATGGAGAT	1512
QY	1914	GATCATGACAGAAAGTATCTCTATAAATGCTTTTCAGTTTCAGATGTTTACTGTTCTACT	1973
DB	1513	GATCAGCAGCAGCAGTGTGGTGACTCTCTACTGTCCAGTTCTTTCACGCACTACT	1572
QY	1974	CTGGCTCGACATCTTATTTGAAGACAGAAATGTTATCTCTGTCATTACTTGAACCTCTGCTA	2033
DB	1573	CTGGCTCGAATGCTCATCAACAGAGAAACCTTGAAGACATTATCATTTAAGACTTTTATG	1632
QY	2034	GAAGTTTTT-----ACCTGAGTACTTGCAGAGGAAACAAATAAATTTCAACTTCCAGGGTTAT	2087
DB	1633	GATCATTTGAGACATCGAGATGCCAGGGCAGATTTTCAGTTTGAACGATACACTGCTTTA	1692
QY	2088	AGCCAGGACAAATTTGGAGAAGTATATGCAAGTATATGTCACCTTAAAGTATATCTCTGATC	2147
DB	1693	CAAGCTTCAAAATTTAGGAGAGTACAGAGCCTTATTTTAGATCTCAAGTATGTGTTAAT	1752
QY	2148	AGCAAAACCAATATGGAAGAAAGATTAAGATGCAAGTTCCCTTGAAGGTTTTCGATCT	2207
DB	1753	AGCAAAACCACTGAATGGTCAGATGAGCTGAGGAGAGATTTCTAGAAAGGGTTTGAATGCC	1812
QY	2208	TTTTTGAAGATTTCTTACTGCTATGCGGAAATGGAAGAAATCCGAAGACAGGTTGGGCAA	2267
DB	1813	TTTTTGGAAATTTACTTAAATTTGTCAGGGATGGATCCAAATTTACAGCTCAAGTAGGACAA	1872
QY	2268	CACATGGAAGTGGATCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAT	2327
DB	1873	CATATTGAAATGGAACAGAGTGGGAGCAGCCTTCACTACATAAAATGAAATTAACAT	1932
QY	2328	ATTTTACTCATGTTTCCAGAGTGGTGTGCTTGTGATGAGAACTCTTACTTTGTGCTTAT	2387
DB	1933	GTCAATTTCAATGATGCAAGGACTGGTGTGCTTTAGATGAAAGGTTTAAATCGAAGCTTAC	1992
QY	2388	AAAGAAATGTCACAAAGCTGTCATGAGGTGCACTACCACTTTTCATATCTAGTAGCAAGACA	2447
DB	1993	AAGAAATGCTCGCTGCTACTGATGCAAGTGTCTATGTTGTTATCTATGATGGTGAACAGCA	2052
QY	2448	GT---AGTCAATCGTGTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTGAGAT	2504
DB	2053	ATCACACTAAGCATTTTGTGGACATTCAGTGGAAACATCAGATACGTGTTTCCCAAGAA	2112
QY	2505	CTTGTAGCATACATCTGCCACTCTTAGGACCCCTTGTGCTCTTCAAGTACGTTTAAAGC	2564
DB	2113	AAAGTAGCATTCCTCCCGAGTTTCTGCTTACTTGCAGGTTTACATGATTTATTTAAGC	2172





Db 4312 ATCCAAAGCATAGAAAGAAATTTTGATGATGAAGATAAAACCAATTTGTTGGTCTTTTACCT 4371  
Qy 4740 AATAGGACGATATATGGTCTGAAGCAATTAATGCAAGTTTGGCAGTGGACAGAGATACC 4799  
Db 4372 TGCAGACTGGATGACTGTCTTAGGTGCAATGACGAGATTTGGCGACGACACTGGACAGTG 4431  
Qy 4800 TGTCTCAGGTCTTGATACAGAAACATCTGGTTCGTCTTCTATCATGATGTTTCTTCTTAAC 4859  
Db 4432 GCATCAGTTTCAGTGGTGAAGGACATTTTGTAAACCTTTTGTGCATCACTGGTGCCTAAT 4491  
Qy 4860 ATAAATCAGAGATACACCATGCTCTGCTCTATAGATCTGTTCATCTATGTTGGTGGT 4919  
Db 4492 GACGCCATGAGGAATCTTCATGTCATATAGATATTGACATGTTTCAATTTATTTGTTGGC 4551  
Qy 4920 GCTGTGTAGCATTCCTCATCTTGTATTTGGATGACCTGTTGATCTGCAGCCTTCTCA 4979  
Db 4552 TTGGTGTGCAATTTCTCGCTGCGTGCAGTG-----TCAGGATTTTTCAGGG 4596  
Qy 4980 GTTAGTCTCTCTATAACCACTTTATCTCTTCCATTTGATCACCATGGCCACACATGCTT 5039  
Db 4597 ATCAGCCTGGGACTGGAGACCTTCACATTTTCCATCTGGTACTATGCGCACATCATATA 4656  
Qy 5040 CAGATPACTTACAGTAGACACAGGCC---TACCCCTTGTCTCAGGTTCAAGAGACAGT 5096  
Db 4657 CAGATCTTACTTACCTCATGTACAGAGAAATGCGATGGATCAAGAAAAATCCCCCTGT 4716  
Qy 5097 GAAGAGGCTCATTCGGCATCTTCTTCTTGGAGAAATTTCTCAATATACAGTGGCTCC 5156  
Db 4717 GAAGAGAAATACCAATCCGCTGGCATCTGTGGAGGAGTGTACAGAGCTGGAAATATGCT 4776  
Qy 5157 ATGGGTGTGATATTCCTGGCTGTGATTTTGGGTCTCACTGAAAGATGGCATCACCCCT 5216  
Db 4777 TTGAAGAAATACCAATCCGCTGGCATCTGTGGAGGAGTGTACAGAGCTGGAAATATGCT 4836  
Qy 5217 TATCTTCCTGTCTGCTGATGTTGTTTCCACTATTTACTTTGGGTAATCCGCTGAGGAA 5276  
Db 4837 TTCTGAGTGTCTGCTTTATTTTTCATTTAATGAGTTCTTCCACCCGAC 4896  
Qy 5277 CTGCATACCAATCTGCGAGAGGAGTACAGTGCATCTGTAGCTATCTATCTTTTACT 5336  
Db 4897 ATTC---AGTTCCTGGAACAGCCATTTTGAACATTTATGTAGCTATCTTCCCTACCA 4953  
Qy 5337 ACATATTTGTTCTCTCTCCAGCAATATTTGGGATCTGTAAAGCCCTTGTCCAGAGG 5396  
Db 4954 AACACCTCATTTGCTTTTCAAGAAATAGTGAGATAATGAATCTGATGAAAGT 5013  
Qy 5397 TGGTGTGAGATCTCTCTTAACTGTTTGAAGCAAAAAACACCGTGTGTCAGGTAC 5456  
Db 5014 TGGTCCGTAAAGTAAAGATATCTAGAAGGTGAAGAGATGCTATAAGATAT 5073  
Qy 5457 CTTAGAAAAAGAAATAGTTGATAGAGCTTCTGTATGATCTAGCTGCTCTGTGATCAA 5516  
Db 5074 CCAAGAGAAATCTAACAAATTAATAAACCCTTCCAGAGGATTAACAGAGCTCATTAATCAA 5133  
Qy 5517 GCTTCTCATTTTCAGGTGCCCCGCTGTCAGATGATGAGCAAAACATCTGTCTCTGCG 5576  
Db 5134 GCATCAAATTTCTGTGCCCCGAAATCAGTGGTGATTAAGACAGAGCCCCAACTCTGTGC 5193  
Qy 5577 CTTTTCTGTGGGCTATATCTATGTTCTCAGAACATTTGCTGCGAGAAATTTGTGAACGG 5636  
Db 5194 CTTGTGTGGGATCTCTGTGTGCTCCAGAGTTACTGTGCCAGCTGAACTGGAAGGG 5253  
Qy 5637 GAAGAGTTGGAGCTTGCAATTTTTCAGGACTTCTGATGATCTAGCTGCTCTGTGATCAA 5696  
Db 5254 GAGGATGTAGGAGCCTGACAGCTCACACCTACTCTGTGCTCTGAGTGGGCACTCTTC 5313  
Qy 5697 CTAAAAATCAGAGATCCGAGTGTCTGTGTTGAAGTTAAAGCCAGAGGCTGTGCTTAT 5756  
Db 5314 CTGAGAGTACGGAAATGTCAGGTGCTATTTTGTAGCTGGCAAAACCAAAGGCTGTTTTAT 5373  
Qy 5757 CCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTTGGCTGAGAGGGGCAACCCC 5816  
Db 5374 TCTCCTCTTACTTGTGATCTATGGGAGACCGACCGGACTCAGACGGGAAATCCT 5433

Qy 5817 CTTCAATTTATCTGCTGAGCGGTATCGAAGCTCCATTTGTTGCTGGCAACACACTGCATT 5876  
Db 5434 TTACATTTATGCAAGAGCGAATTCAGAGATTCAGAGCTCTGGCACCAACACAGTGTCTC 5493  
Qy 5877 ATAGAAGAGATTTGCTAGGAGCCCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCA 5935  
Db 5494 ACAGAGAAATTTGGACATGCACAGGAAGCCAAATCAGACACTGGTTGGCATTTGACTGGCA 5552

## RESULT 10

US-10-758-636A-3  
; Sequence 3, Application US/10758636A  
; Publication No. US20050089876A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN B3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966C  
; CURRENT APPLICATION NUMBER: US/10758,636A  
; PRIOR FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 6300  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (295)..(5559)  
US-10-758-636A-3

Query Match 21.9%; Score 1381.8; DB 21; Length 6300;  
Best Local Similarity 56.5%; Pred. No. 0;  
Matches 2947; Conservative 0; Mismatches 2167; Indels 105; Gaps 16;

Qy 780 TGGGATCAGCAGTTGATTTTATATCTGCTTTCTTGATCATTTTGGCAATTTGGTGCCA 839  
Db 376 TGGTGCAGCAACTGACCTCACTAGAGAGTGTACCAGCACTTTAGCCCACTATGTATCCC 435  
Qy 840 GAAATTTACTTTGCTGAAATGCGACCCAGACTTGGAAAGAGGAGGAAAGTGTACAAATG 899  
Db 436 AAAATCTACTCGAGGGTCCCAACCTTTTCCACGAAAGAGACATGCTGGCAGCAT 495  
Qy 900 TCAATATTCACTCCCACTGGATGGTACTTATTTGGAGAAGATCCAGATATTTCCTTAGAG 959  
Db 496 GTTTTGTGGACCAATGGATGGTACTCTTTGTGTGAAGATCTTGCAATTTGGATTTCCA 555  
Qy 960 AATTTGAAGCAC---AGTGGAGCAATTTAGCTTTTGGAGGGTTTCAAAGTGGAGAG 1016  
Db 556 AAATCTGAGCAAGCAAAACAACTCTCATCTTTGTGTGCTGTGTTTAAAGTAGGAGAG 615  
Qy 1017 ACAACCTATTCTTCAGGGATTTGCAATTCATCCAACTGTGTACTCTGTATGAGCTGC 1076  
Db 616 CTTACATATTCTTGACAGACTGTGCGTTGATCCAACTGTGTGTTTGTGATGGAGTGC 675  
Qy 1077 TTCAGGACAGTGTTCATAAAATCATCGTTTACAGATGCAATCTTCTACCTGGAGGGG 1136  
Db 676 TTTTGGGAGTATTCACAGAGATCATCGATATAGGATGACAAACATCAGGAGGTGGAGGT 735  
Qy 1137 TTCTGTGATCTGTGAGACACAGAGCATGGAATAATGCGCCCTTTTGTGTAAATCATGAA 1196  
Db 736 TTCTGTGATCTGTGTGATCTACTGAAAGCCCTGGAAGAGGGTCTTACTGTGCAAAAACATGAA 795  
Qy 1197 CCTGGAAGCAGGTACT---ATAAAGAGAAATTCAGCTGTCCGTTGCAATGAAGAGTGA 1253  
Db 796 CTTAACACCTCTGAATTTGAGGAAGAGAGATCTCTTGTTCATTTATCAGAGATGTG 855  
Qy 1254 ATTGTCCAGCCAGGAAATATTTCTTTCAGTGATGATAAAATATGTGCTAGAAATGACTATA 1313

Db 856 ATAGCAAGAACTTATAACATTTTCTATTAGTTTCGGTATGCGTAGAATAATTAAACC 915  
Qy 1314 TGGGAAGAGGAAAGAAAGAACTGCTCTGAACTCCAGATAAGGGAGAGAAATAAGAAAGTAC 1373  
Db 916 TGGGAAAGAAAGAAAGTGAATTCAGCAGAGATTTAGAGATGGTAGAAGAGTGACACCTAC 975  
Qy 1374 TATTGTGTCCTTTTCAATGATGAACACCAATTCATATGACCAGCTCATATACAGCCTACAA 1433  
Db 976 TATTGCATGCTGTTTAATGATGAGGTTTCAACCTATGAAACAAAGTATTATTATCTTTCAG 1035  
Qy 1434 AGAGCTCTTGAGCTGAGCTGCGAGAGCCAGTTGCGATACCACTGCCATTGACAAAGAG 1493  
Db 1036 AAAGCTGTTAACTGTACACAAAAGAAAGCTATTGGTTTTGCACTACAGTAGATCGAGT 1095  
Qy 1494 GGTGCTGGGCTGTTAAAGCGGAGCTTATGCTGCTGCGAGAAAGCAAAAGAAAGATATA 1553  
Db 1096 GGGCGTAGCTCTGTTGCGATATGGAGATTTTCAAGTATGTGAGCAAGCAAAATCAGTAAT 1155  
Qy 1554 AAGAGTCATTGAGAAATGCTCTCAACATCCACTTCATGTAGAGTATTACACTCAGAG 1613  
Db 1156 GTGAGAAATACAGTAGACAGAC-----AAAGCCACTCAAAGTTTCAAGTTATGCTTCGTCT 1212  
Qy 1614 ATTATGGCTCATCAGAAATTTGCTTTGGCTCTTGGTCTCTGATGAAACAAAATATTAGAGC 1673  
Db 1213 ATTGTCGCACATCAGAAATTTGGTTTGAACCTTTTGTCTTGGCTGGGAGTATTATTGGA 1272  
Qy 1674 TATTCAAGTGACTTTAGGCAATCTTTTGCAGCATGCTTTAGAGAGAAAGCACTGACTCG 1733  
Db 1273 TATTCAAGATGCTTTCGCGGATTTTATGTCAAGTTGGTTTACAAGAGGCGCAGATGGT 1332  
Qy 1734 GAGATCCCTGCTCATTAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCC 1793  
Db 1333 GAAAACTCTTCTAGTCGACAGACTGATGCTTAGTGATTCGAAATTAAGGAGGTTGCT 1392  
Qy 1794 CGTAAGATCTTTCATGAATGCACTTTCAGCAGTTTATTTATGGAGATGGAATACAAAAA 1853  
Db 1393 AGGAGTGATATCATCAGTTGTTTATGAGCAGTCTGCTTATGATTTGAAATACAGAAA 1452  
Qy 1854 CTCTTTGCTATGGAATTTGGAAGTATTATAAACAAGCTGCGAGAAAGAAATATATCAGTAT 1913  
Db 1453 CTATTTGCTGTTGCAATTTTCAAAAAATTTACAGCAGTTTGCAGAGAGATTTTATGGAGAT 1512  
Qy 1914 GATCATGACAGAGTATCTTATTAAGTATGCACTTTTCAAGTTTCAAGTTTATGCTTACT 1973  
Db 1513 GATCAGGAGCAGAGTGTGCGTGAATCTCTATCTGTCCAGTTTCTTCAAGCAGCTACT 1572  
Qy 1974 CTGCTGACATCTTATTGAGAGCAGAAATGTTATCTCTGTCATTTACTGAACTCTGCTA 2033  
Db 1573 CTGCTGATGCTCATCACAGAGAAACTTTGATGAGCATTTATCAATTAAGACATTTATG 1632  
Qy 2034 GAAAGTTTT-----ACCTGAGTACTTGGACAGGAAACAAATAATTCACCTTCCAGGGTTAT 2087  
Db 1633 GATCATTTGACATCGAGATGCCAGGGCAGATTTCAAGTTTGAACGATACACTGCTTTA 1692  
Qy 2088 AGCCAGGACAAATTTGGGAAGATATATGCAATATATGATGACCTTAAAGTATATCTGATC 2147  
Db 1693 CAAGCCTTTCAAATTTAGGAGAGTACAGAGCCTTATTTTAGATCTCAAGTATGCTTAAT 1752  
Qy 2148 AGCAAAACCAATATGACAGAAAGATTAAGATGCAATGCTTCTTGAAGTTTTCGATCT 2207  
Db 1753 AGCAAAACCAATGATGCTGAGTACAGTACAGTACAGGAGATTTCTTAGAAGGTTTGAATGCC 1812  
Qy 2208 TTTTGAAGATTTCTTACCTGTATGACAGGAAATGGAAGAAATCCGAAGACAGGTTGGGCAA 2267  
Db 1813 TTTTGGAAATTAATAATGATGACGGAATGATCCAAATTAACGTCAGTAGGACAA 1872  
Qy 2268 CACATTAAGTGAATCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTAAGAAAT 2327  
Db 1873 CATATTGAATGGAACCAAGAGTGGAGAGCCTTCACTACAAATGAATTAACACAT 1932  
Qy 2328 ATTTTACTCATGTTCCAGAGTGTGCTTGTGATGAGAACTCTTACTTTGCTGCTTAT 2387  
Db 1933 GTCAATTTCAATGATGCGAGACTGGTGTGCTTTTCAATGAAAGTGTTAATCGAAGCTTAC 1992

Qy 2388 AAAGAAATGTCACAAAGCTGTGATGAGGTGCGAGTACCAAGTTTCATATCTAGTAGCAAGACA 2447  
Db 1993 AAGAAATGCTCGCTGTACTGATGCGAGTGTCAATGCTGTGTATCTGATGGTGAACAGCCA 2052  
Qy 2448 GT---AGTACAACTCGTGTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTGAGAT 2504  
Db 2053 ATCACACTAAGCAATTTTGGACATTCAGTGGAAACTATCAGATATCTGTGTTTCCCAAGAA 2112  
Qy 2505 CTTTGAAGCATACATCTGCCACTCTCTAGGACCCCTTGTGCTTCTCATGTAGCTTTAAGC 2564  
Db 2113 AAAGTTAGCAATTCACCTCCAGTTTCTGCTTACTTTGAGGTTTACATGATATTAAAGC 2172  
Qy 2565 AGGCTGGGTGCTGTTTCAAGACTGCAATGAAATTTGTGCTCTTTTGAGGACTTTTCAAGTAGAG 2624  
Db 2173 AAAAGTGAAGTGGCATATAAATTTCCAGAGCTCTCACTCTAAGTGAACCTTAGCCCAACC 2232  
Qy 2625 GTACTAGTGAATATCCTTTTACGTTTGTGCTGTGTTGTTGCTGCGAGGTTTGTGCTGAGATG 2684  
Db 2233 ATGTTGATAGAACACCTCTTTAGATGCTTGTCTGTGCTGCCAAGTACATGCCCGAAATG 2292  
Qy 2685 TGGCGAAGAAATGGAATGCTCTCTTATTAGCCAGGTTTATTATTACCAAGATGTTTAAAGTC 2744  
Db 2293 TGGAGAGAAATGGTCTCTCTAGTAAACCCAGATTTATTACTACCAATATGTGAATGTC 2352  
Qy 2745 AGAGAAAGAAATGATGATGATAAAGATATCATCATGCTTCAAGATTTGTTGCAATCTTTAATGAT 2804  
Db 2353 AGAGCTGAGATGTTTGAACAGGATGTAGTAATGCTTTCAGACAGGTTCTCCATGATGAT 2412  
Qy 2805 CCCAATAGTCTTGTACTGTTACTGTTACTCAGAGGATGAACTTCCGAGGCTTTTAAACAAG 2864  
Db 2413 CCNAATCATTTTCTGATGATCATGCTCAGCCGCTTTTGAACCTTTATCAGATTTTCACTACT 2472  
Qy 2865 ACCATATCTACAAA-----GACCAGGATTTTGAATTTAAACAATAT 2903  
Db 2473 CCAGACTATGGAAGAAAGATTTAGTTCTGAGATTACCCATAGGATTTGTTTACGCGAAGC 2532  
Qy 2904 AATACACTAATPAGAAGAAATGCTTCAAGTCTCATCTATATTTGTTGGGTGAGCGTTATGTA 2963  
Db 2533 AATACTCTAATAGAAGAAATGCTATACCTCATTAATAATGCTTGTGGAGAGAGATTTAGT 2592  
Qy 2964 CTTGAGTGGGAATGTGACCAAGAGAGGTGCAATAGAGAAATCATTTCACTTGTCTT 3023  
Db 2593 CTTGGAGTTGGAAGAGGTAATGCTACAGATGAAATCAAGCGAGAGATTTATCCATCAGTTG 2652  
Qy 3024 TGCAATGAAACCAATGCGACACAGTGCCATTCGCAAAATTTTACCTGAGAAATGAAATAT 3083  
Db 2653 AGTATCAAGCTTATGGCTCATAGTGAATTTGTAAGTCTTTTACCTGAGATGAGAAACAG 2712  
Qy 3084 GAAACTGGCTTAGAGAAATGTCATAAAACAAAGTGGCCACATTTTAAAGAAACCAAGGTGTATCA 3143  
Db 2713 GAGACTGGCATGGAGAGTGTAAATCGAAGCAGTTGCCCATTTTCAAGAAACCTGGATTAACA 2772  
Qy 3144 GGCCATGGAATTTATGAATTAAGAGATGAATCACTGAAAGACTTCAATATGATCTTTTAT 3203  
Db 2773 GGACGAGGCAATGATGAATGAAACCAAGATGTGCCAAAGAGTTTCACTGTTTCTTAT 2832  
Qy 3204 CATTACTCCAAAACCCAGCATTAGCAAGCTGAAACATATGCAAGAAAGAAAGAGAAACAA 3263  
Db 2833 CACTTTTCAAGGGCAGAACAGTCCAGGAGAGAGGCGCAACCGAAATTTGAAAGACAA 2892  
Qy 3264 GAAACAAAGATGAAGCAATTTGCCCCACCACTCTCTGAAATTTTGCCTTGTCTTTCAGC 3323  
Db 2893 AATAGAGAGATACAGCACTCCCACTCCGGTGTGGCTTCCCTTCTGCTTCTGTTTGA 2952  
Qy 3324 AAAGTGAATTAACCTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAG 3383  
Db 2953 AGCTGTTTAACTTTTTCAGTTCAGATGTCATGTTGTGTCATCATGGAACAAATTTCTGCAA 3012  
Qy 3384 CGGCAATAGACACAGATTTCTAACTTGTGGACCGAAGGGATGCTCCAAATGCTTTTTCAT 3443  
Db 3013 TGGGCTGTGGAACATAATAGGATATGCTTGGTTCAGATCCATGCTGCAAGGGTGTATCAT 3072



Db 5194 CTTGTGTGGGATCTCTGCTGTGCTCCAGAGTTACTGCTGCCAGCTGAACCTGAAGGG 5253  
Qy 5637 GAAGAGTTGGAGCTTGCATTTTTCACGCACTTCTCACTGTGGAGCCGAGTCTGCAATTTTC 5696  
Db 5254 GAGGATGTAGGAGCTTCACAGCTCACACCTACTCTCTGTGGCTCTGGAGTGGGCATCTTC 5313  
Qy 5697 CTAAATATCAGAGTCCGAGTGTCTCTGTGAGGTGAAGCTTAAGCCAGAGGCTGTCCCTAT 5756  
Db 5314 CTGAGGTACGGGAATGTCAAGTGTCTATTTTGTGTCGCAAAACCAAGGCTGTTTTAT 5373  
Qy 5757 CCAGCTCTCTTACTTGGATGAATATATGAGAAACAGACCTTGGCTCAAGAGGGGCAACCC 5816  
Db 5374 TCTCTCTTACTTGTATGACTATGTGGAGACCGACCGAGTCTCAGACGGGGAATCT 5433  
Qy 5817 CTTCAATTTATCTGTGAGCGGTATCGGAAGCTCCATTTTGGTCTGCAACAACACTTGCATT 5876  
Db 5434 TTACATTTATGCAAGAGCGATTCAAGAAGATTCAAGAAGCTCTGGCACCACCAACACAGTGT 5493  
Qy 5877 ATAGAGAGATTGCTAGGAGCCCAAGAGACTAATCAGATGTTATTTGGATTCAACTGGCA 5935  
Db 5494 ACAGAGAAATTTGGACATGCAAGAGGCAATCAGACACTGGTTGGCAATGACTGGCA 5552

## RESULT 11

US-10-758-672A-5  
; Sequence 5, Application US/10758672A  
; Publication No. US20040185037A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966B  
; CURRENT APPLICATION NUMBER: US/10/758,672A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 5  
; LENGTH: 6089  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (766)..(6030)  
US-10-758-672A-5

Query Match 19.6%; Score 1233.6; DB 19; Length 6089;  
Best Local Similarity 54.7%; Pred. No. 0;  
Matches 2870; Conservative 0; Mismatches 2269; Indels 105; Gaps 17;  
Qy 769 TGGCATCTTGTGGATCAGCAAGTTGATTTTATCTGCTTCTTGATCATTTGGCAC 828  
Db 836 TGCAGGAGATGGCTGCAAGCAACCGACCTCAACAGAGAGTGTACAGCAATTTAGCCC 895  
Qy 829 AATTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGAGGAAA 888  
Db 896 ACTGTGCCCCAAATATCTACTGCGGGGCCCTTAACTTCTTCCCTCAGAGAGAGACACGC 955  
Qy 889 GTGTACAAATGTCAATATTTCACTCCACTGGAAATGGTACTTATTTGGAGAAGATCCAGATA 948  
Db 956 TGGCAGACACATCTCTGTGGGACCGATGGAGTGTATCATCTGCGCTGAAGACCTTGGC 1015  
Qy 949 TTTGCTTAGAGAAATTTGAAGC---AGTGGAGCAATTTAGCTTTGTGGAGGGTTTCA 1005  
Db 1016 TGGGATTTCCAAAGCTCGAGGAGGCAAAACAAGCTTCTCACCTCTGTGGCCGAGTGTTA 1075  
Qy 1006 AAAGTGGAGAGCAACCTTATCTTCCAGGATTTGTCAATTTGATGATCCAACTGTGTACTCT 1065  
Db 1076 AAGTGGGGAGCACTACATCTCTCGAGAGTGTGCAAGTGTGACCCACCTGTGTAT 1135  
Qy 1066 GTATGGACTGCTTCCAGGACAGTGTTCATAAAATCATCGTTACAAGATGCATCTTCTA 1125

Db 1136 GATGGAGTGTCTCTCGGAAAGTATCCATAGAGACCATCGATATAGGATGACCAATCGG 1195  
Qy 1126 CTGAGAGAGGGTCTGTGACTGTGAGACACAGAGGCACTGGAAGACTGGCCCTTTTGTG 1185  
Db 1196 GAGGAGGGGCTCTGTGACTGTGAGACACAGAGGCGTGGAAAGAGGACCTTACTGCC 1255  
Qy 1186 TAAATCATGAACCTGGGAAG---AGCAGGTACTATAAAAGAGAAATTCACGCTGTCCGTGA 1242  
Db 1256 AGAAGACACAGCTCAGCAGCTCTGAAGTTGTGGAGGAGGATCCTCTTGTGCATCTAT 1315  
Qy 1243 ATGAGAGGTAAATGTCCCAAGCCAGGAAATATTTCTTCACTGATATAAAATATGCTAG 1302  
Db 1316 CAGAAGATGTATCGCCAGAACTTACAACATTTTGTCTATTTATGTTTCGATATGCAATG 1375  
Qy 1303 AAATGACTATATGGGAAGAGGAAAGAACTGCTCTCTGAACTCCAGATAAAGGGAGAAA 1362  
Db 1376 ATATACTGACCTGGGAAAGAAAGTGAATGCTTGAAGACTTGAAGTGGCAGAGAGA 1435  
Qy 1363 ATGAAGATATCTATTTGTGCTCTTTTCAATGATGAACACATTCATATGACACAGTCATAT 1422  
Db 1436 GTGACACCTACTACTGCACTGCTGTTTATATGATGAGTTTCAACCTATGAGCAAGTCAAT 1495  
Qy 1423 ACAGCTACAAAGAGCTCTTCACTGTGAGCTGCGAGAGCCAGTTGTCATACCACTGCCA 1482  
Db 1496 ATACCTCTCAGAAAGCTGTGAACCTGTACAGAGGAAAGCAATTTGGCTTTGCACTAC 1555  
Qy 1483 TTGACAAAGGGTCTGCGGCTGTAAAGCGGAGCTTATGCTGCTTTCAGGAAAGCA 1542  
Db 1556 TTGATCGAGTGGCGGTAGGCTGTCCGATATGAGATTTCCAGTACTGTGATCAAGCA 1615  
Qy 1543 AGGAAGATATAAGAGTCAATTCAGAAATGCTCTCAACATCCACTTTCATGTAGAAGTAT 1602  
Db 1616 AGACAGTCAATGTGAGGAAACACAGCAGACAGACCAAG---CCGCTCAAAGTTCAAGTTA 1672  
Qy 1603 TACACTCAGAGATATATGCTCATCAGAAATTTGCTTTCGGTCTTGGTCTCTGGATGAACA 1662  
Db 1673 TGCACCTCTCGTGGCTGCTCATCAGAAATTTGGTTTGAAGCTCTGCTGGCTGGGAA 1732  
Qy 1663 AAATATGAGCTATTTCAAGTCACTTTAGGCAGATCTTTTGGCAAGCTCTTTAGAGAG 1722  
Db 1733 GTGTTATTTGGATATCTCAGATGGCTTTCGAGGATTTTGTGTCAGTTGATTAAGAG 1792  
Qy 1723 AACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTAAATGCTTTGGGATGCAAGCTTT 1782  
Db 1793 GTCCAGATGGCGAAAATCTCTCTCTGTCGACAGACTGATGCTTAATGATTTCAAATTTAT 1852  
Qy 1783 ATAAAGTGGCCGTAAGATCTTCAATGAATGATCTTCAAGCTTCTCAGCAGTCTTTTATGGAGTGG 1842  
Db 1853 GGAAGGGGCTTAGGAGTGTGTATCACCAAGTTGTTTCAAGAGAGCTGCTCATGAGCCTCA 1912  
Qy 1843 AATCAAAAAAATCTTTGCTATGGAATTTTGAAGTATTTATAAAACAACCTGCAGAGAAAGAT 1902  
Db 1913 AGTATAGAGCTGTTCGCGCTTCGATTTGCTAAATTTACCGCAGTTGCAGAGGATTT 1972  
Qy 1903 ATATCAGTGAATCATGACAGAAAGTATCTCTATAAATGCACTTTTCACTTTCAGATGTTTA 1962  
Db 1973 TTATGAGGATGATCAAGAGCGGCGAGTGGTGAAGTCTGCTCTGCTCTCAGTTCTTCA 2032  
Qy 1963 CTGTTCTTACTCTGCTGCAATCTTATTTGAAGAGCAGAAATGTTATCTCTGTCAATCTG 2022  
Db 2033 CCGCACCGACGCTGGCGGATGCTCTCTCAAGAGAGAACTGTGATGACCGTTCATTA 2092  
Qy 2023 AAACTCTGCTAGAGTTT-----ACCTGAGTACTTGGCAGGAAACAATAAATTTCAACT 2076  
Db 2093 AGGCTTTTATGAGCAATTTTGAACAACAGAGATGCCCGAGGAGATTCAGTTTGAACGCT 2152  
Qy 2077 TCCAGGGTTATAGCCAGGACAAATTTGGAGAGATATATGCAAGTATATGTCACCTTAAAGT 2136  
Db 2153 ACATGCGCTTCAAGCCTTCAAGTTTCAAGGATCCAGAGCTCATCTTAGATCTCAAGT 2212  
Qy 2137 ATATCTGATCAGCAAAACCCCAATATATGACAGAAAGATTAAGAAATGAGTTCCTTGAAG 2196





QY 4309 ATACTGTGATCCCATTTATTCCTTTGCAACCTCAAAAGATATAACAGTGAGATGAGATG 4368  
DB 4418 ACACGGTGTATCCCC-----CTGCTGCTTCTCCAGGAGCATCTCCAGCAGGAGTTAA 4471  
QY 4369 CTCCTGTCTCAACTTTTGTGACCTTGGCAGCGGTGATACAGACTGTCTGGCCAGAAATATCAG 4428  
DB 4472 ATTTTTCAGACCAACAGATCTGGCACAAGTGACAGTGACGAGAGCATACACAGCATTAAGG 4531  
QY 4439 GTTATAATATAAGACATGCTAAGAGAGAAACCCAAATTCCTATTTTCTTTAATCAAGAA 4488  
DB 4532 TGGTCCAGATGTGAGGAGAAAGCAATGCTGCTGACACAGCTCTTCTTCAGAGGACACAG 4591  
QY 4489 TGGGAGATTTACTTTTGGAGTTCCATTCATCTGAGTTTGGCGTTGAGTCTTCGATTA 4548  
DB 4592 AAGCCATGAATATAATACCGATCCCGAAGGCTTCAGGCGCTGATTTTATCCTAGGAACC 4651  
QY 4549 AATATTTCAAATAGCATCAAGGAATGTTATTCTCTTTGCCACAACAATTTATAGAAATG 4608  
DB 4652 CATATTTCTGATAGCATAAAGGAATGTTAAACGACATTTGGAACGCTGCTTACAGGTGG 4711  
QY 4609 GATTGAAGTGCACCTGATGAAGAGGATCTCTGAGTCCCCATCTGACCTGGAGCACCT 4668  
DB 4712 GACTGAAGGTTTCATCTTAATGAAGTGACCCCGTGTGCCATCTGCTGTGGGGGACCT 4771  
QY 4669 GCGCTTCACTATCCAGGCAATGAATCTATTGGGAGATCAAGGAACCTCTGTTG 4728  
DB 4772 GTGCATACACCATCCAGAGCATAGAAAGAAATTTTGAGTGAGAGAGGCGCTGTTTTG 4831  
QY 4729 GAGCACTTCAAAATAGGAGCATATGCTGCAAGCAATTAATGAGCTTTGAGTTGCAC 4788  
DB 4832 GACCTCTGCTTGTAGACTGAGAGCTGTCTCAGGTGTTAAACGTTTGCAGGACAC 4891  
QY 4789 AGAGGATTAACCTGCTCCTCAGTCTCGATACAGAAACATCTGTTGCTCTTCTATCAGTTG 4848  
DB 4892 ATTGCACAGTGGCGTTACTTCTGCTGTTAGAGGACACTTCTGTAAACTCTTTGTCATCT 4951  
QY 4849 TTCTTCTTAACATAAAATCAGAGATACACCATGCTCTGCTCTATAGATCTGTTTCAATG 4908  
DB 4952 TGGTGGCTAGTGACAGCTATGAAGACCTCCCGTGCATACAGACATGACATGTTTCACT 5011  
QY 4909 TTTTGGTGGGTGCTGTGTTAGCATTCCTATCCCATCTCTGTTATTTGGATGACCTGTTGATGCG 4968  
DB 5012 TGCTGGTGGGCTGTGCTGTCTTTCCAGCTCTGAGTG-----TCAGG 5056  
QY 4969 AGCCTTCTTCAGTTAGTTCTTCTCTATAACACACTTTATCTCTTCCATTTGATCACCATGG 5028  
DB 5057 ATTTTTCAGAGAGCAGCTGGCCACTGGGACCTGCACATCTTCCACTTGGTTACCATGG 5116  
QY 5029 CACATGCTTCAGATATCTATTACAGTAGACACAGGCGCTA---CCCTGTGCTCAGGTTTC 5085  
DB 5117 CACATCGTACAGATCTTACTTACTCTATGTACAGAGAGAAATGGCATGGATCAAGAGA 5176  
QY 5086 AAGAGACAGTGAAGAGCTCATTCGGCATCTTCTTTCTTTGCGAATATTTCTCAATATA 5145  
DB 5177 ATCCCACTGGGGAAGAGAACTGGCCATCTCTTTGACAAAACACTTCAACAGTATA 5236  
QY 5146 CAAAGTGGCTCAATTTGGGTGTATATCTGCTGCTGTTATTTGGGTCTCACTGAAGAAATG 5205  
DB 5237 CTGGAAGTGCCTTGAAGAGAGCCCTCCGCTGACACTGTGGAGGAGCGTCCGGGCGG 5296  
QY 5206 GATCAACCCCTTATCTTCGCTGTGCTGATGTTTTCATATTTTACATTTTGGGGTAACTC 5265  
DB 5297 CCATCATGCTTTCTCAAGTGTCTGCTTGTGTTTTCACATATTTTAAATGGAGTCCCGG 5356  
QY 5266 GCGCTGAGGAAGTACATACCAATTTCTGAGAGGAGATACAGTGCATCTGTAGCTATC 5325  
DB 5357 CCCCTCCAGACCTTCA---AGTTTCTGGAACAAGCCATTTTGAACAATTTATGTAACTACC 5413  
QY 5326 TATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATATCTGTAAAGCCCT 5385  
DB 5414 TTTCCCTACCAACCACTCTATCACTTTTCAAGAAACAGTGCATCATGAACTCCC 5473  
QY 5386 TGCTCCAGAGGTGTGTGAGATCCTGCTTACTAACTGTTTGAAGCAAAAAACACCG 5445

DB 5474 TGAATTGAAGTTGGTGGCCAGAACAGTGAAGTTAAACGGTATCTAATCGCAGAGAGGAG 5533  
QY 5446 TGGTCAGGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTGATGACTATAGCTGCC 5505  
DB 5534 CGATAAGCTACCCACAGAGGAGCTAACAACTGTATAGACTTCCAGAGGATTTACAGAGCC 5593  
QY 5506 TCCTGAATCAAGCTTCTCATTTTCAGGTGCCACCGGTCTGCGAGATGATGAGCGAAAGCATC 5565  
DB 5594 TCATTAACCAAGCATCCAATTTTCTGTCGCCCAATCAAGGTGGCCGACAGAGCAGAGCTC 5653  
QY 5566 CTGTCCTGCTGCTTTTCTGTGGGGCTATCTATGTTCTCAGAACATTTGCTGCCAGGAAA 5625  
DB 5654 CTACTCTGTGCTCTGTGTGGAGTCTCTCTCTCTCAGAGTTACTGTGCTCCCAAGCTG 5713  
QY 5626 TTGTGAACGGGGAAGAGGTTGGAGCTTGCAATTTTTCACGCACTTTCATCTGTGGAGCGGAG 5685  
DB 5714 AGCTGGAGGTGAGAGCGTCCGAGCCTGCACAGCACACACCTACTCTCGGCTCCGGGG 5773  
QY 5686 TCTGATTTTCTTAAATTCAGAGATGCCAGGTGCTCTGGTTGAAGGTAAAGCCAGAG 5745  
DB 5774 CCGGCATCTTCTCAGAGGTGGGGAATGTCAAGTGTCTATTTTGTAGCTGCCAAAACCAAG 5833  
QY 5746 GCTGTGCTATCCAGCTCTTACTTGGATGAATATGGAGAACAGACCTTGCCTGAAGA 5805  
DB 5834 GATGTTTATTTCTCTCTTACCTTGACGACTATGGAGAGACCGACCGGACTCAGAC 5893  
QY 5806 GGGGCAACCCCTTCATTTATCTGTGAGCGGTATCGGAAGCTCCATTTGCTTGGCAAC 5865  
DB 5894 GAGGAATCTTTTACATTTATGCCAGAGCGGTTTCGAAGATCCAGAGCTCTGGCAGC 5953  
QY 5866 AACACTGCTATAGAAGAGATGCTAGAGCCAGAGACTAATCAGATGTTATTTGGAT 5925  
DB 5954 AGCATAGTATCAGAGAGGAGATCGGACACGCGCAGGAGGCTAAACAGACCTCTGTCGAA 6013  
QY 5926 TCAACTGGCAGTTACTGTGAGCTC 5949  
DB 6014 TTGACTGGCAGCATTTTATAATCGC 6037

## RESULT 12

US-10-758-636A-5  
; Sequence 5, Application US/10758636A  
; Publication No. US20050089876A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966C  
; CURRENT APPLICATION NUMBER: US/10758.636A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 5  
; LENGTH: 6089  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (766)..(6030)  
US-10-758-636A-5

Query Match 19.6%; Score 1233.6; DB 21; Length 6089;  
Best Local Similarity 54.7%; Pred. No. 0;  
Matches 2870; Conservative 0; Mismatches 2269; Indels 105; Gaps 17;

QY 769 TGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGATCATTTGGCAC 828  
DB 836 TCGCAGGAGGATGGCTGCAAGCAACCGACCTCAACAGAGAAAGTGATACAGCATTTAGCCC 895

829 AATTGGTGCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGACGAGGAAA 888  
896 ACTGTGTGCCAAATCTACTCCGGGGCCCTAAACCCCTTCCCTCAGAAAGACACGC 955  
899 GTGTACAAATGTCAATATTTCACTCCACTGGAAATGTACTTATTTTGGAGAAAGATCCAGATA 948  
956 TGGCACAGACATCTCTGTCTGGGACCGATGGAGTGGTACATCTGCGCTGAAGACCCCTGGC 1015  
949 TTTGCTTAGAGAAATTTGAAGCAC---AGTGGAGCAATTTTCAGCTTTTGGAGGGTTTCA 1005  
1016 TGGGATTTCCAAAGCTCGAGAGGCAAAACAGCCCTTCTCACCTCTGTGGCCGAGTGTTA 1075  
1006 AAAGTGGAGAGCAACCTATTCTTCAGGGATTTGTGCAATTGATCCAACTGTGTACTCT 1065  
1076 AAGTGGGGAACCTACATACTCTCTGAGAGACTGTGCAAGTTGACCCACCTGTGTGTTTTAT 1135  
1066 GTATGGAAGCTGCTTCCAGGACAGTGTTCATAAAATCATCGTTPACAAAGATGCATCTTCTA 1125  
1136 GCATGGAGTGTCTCTGGGAGTATCCATAGAGACCATGATATAGGATGACCAATCGG 1195  
1126 CTGGAGGAGGGTTCTGTGACTGTGGAGACACAGAGGCAATGGAAACTGGCCCTTTTGTG 1185  
1196 GAGGAGGGGGCTTCTGTGACTGTGTGACACTGTGAGCGGTGGAAAGAGGACCTTACTGCC 1255  
1186 TAAATCATGAACCTCGAAG---AGCAGGTACTATAAAAGAGAAATTCAGCTGTCCGTTGA 1242  
1256 AGAAGCAAGCTCAGAGCTCTGAAGTTGTGGAGAGAGGATCTCTTGTGCATCTAT 1315  
1243 ATGAAGAGGTAAATTTGCCAAGCCAGGAAATATTTCTTCAGTGATGATAAATATGTCGTAG 1302  
1316 CAGAGATGTGATGCCAGAACTTACAACTTTTGTCTATTATTTGCTATATGATGCACTAT 1375  
1303 AAATGACTATATGGGAAGGAGGAAAGAACTGCTCTGAACTCCAGATAGGAGGAGAAA 1362  
1376 ATATACTGACCTGGGAAAGAAAGTGAATTCCTGAAAGACTTAGAAGTGGCAGAGAGA 1435  
1363 ATGAAGATACTATTTGTGTCCTTTCAATGATGAACACCAATTCATATGACACGTCATAT 1422  
1436 GTGACACCTACTACTGTGCTGTTTAAATGATGAGGTTCAACCTATGAGCAAGTCATTT 1495  
1423 ACAGCTACAAAGAGCTCTTGAAGTGTGAGCTGCGAGAGCCAGTTGCAATACCCTGCCA 1482  
1496 ATACCTTTCAGAAAGCTGTGAAGCTGTACACAGAGAGGACCAATTTGGCTTTGCAACTACAG 1555  
1483 TTGACAAAGAGGGTCTGTGGGCTGTATTAAGGGGAGCTTATGCTGCTTCCAGGAGCA 1542  
1556 TTGATCGAGATGGCGGTAGGCGCTGTCCGATATGGAGATTTCCAGTACTGTGATCAAGCA 1615  
1543 AGGAAGATATAAGAGTCAATTCAGAAATGTCTCTCAACATCCACTTCATGTAGAAAGTAT 1602  
1616 AGCAGTCAATTTGGAGGAACAACAGCAGACAGACCAAG---CCGCTCAAGTTCAAGTTA 1672  
1603 TACACTCAGAGATATATGGCTCATCAGAAATTTGCTTTGGCTCTTGGTTCTCGGATGAACA 1662  
1673 TGCACTCTCTCGGTGCTGCTCATCAGAAATTTTGGTTTGAAGCTCTGTGCTGGCTGGGA 1732  
1663 AAATATGAGCTATTAAGTGAATTTAGGAGATCTTTTGGCAAGCATGCTTTAGAGAG 1722  
1733 GTGTTATTTGGATACTCAGATGGCTTTCGAGGATTTTGTGTCAAGTTTGGATTTACAAAG 1792  
1723 AACCTGACTCGAGAAATCCCTGCTCTCATAGCAGGTAAATGCTTTGGGATGCAAGCTTT 1782  
1793 GTCCAGATGGGAAACTCTTCTCTGTGCGACAGACTGATGCTTAATGATTTCAAATAT 1852  
1783 ATAAAGGTGCCGTGAAGTCTTTCATGAATTCATCTTCAGCAGTTTATTTTATGGAGATGG 1842  
1853 GGAAGGGGCTAGGAGTGTATCACCACTGTCTTCAAGCAGGCTGCTCATGGACCTCA 1912  
1843 AATACAAAAATCTTTTGTCTATGGAAATTTGTGAAGTATTTATTAACAACTGCGAGAAAGAT 1902  
1913 AGTAAAGAAAGCTTTTCGGCTTCGATTTTGTCTTAAATTTTACCGGAGCTTTCAGAGGGAT 1972  
1903 ATATCAGTGTATCATGACAGAGTATCTCTATAACTGCACTTTTCAGTTTCAGATGTTTA 1962

1973 TTATGGAGATGATCAAGAGCGGCGAGTGTGGTGACTGCTCTGCTGTCTGTCAGTTCTTCA 2032  
1963 CTGTTCTTACTCTGCTCGACATCTTATTTAGAGAGCAGAAATGTTATCTCTGTCATTACTG 2022  
2033 CCGACCCGAGCTGTGGCGAATGCTCTCTCAGAGAGAACTCTGATGACCGTTATCATTA 2092  
2023 AAACTCTGCTAGAAAGTTTT-----ACTGTAGTACTTTGGACAGGAAACAATAAATTTCACT 2076  
2093 AGGCTTTTCATGGACATTTTGAACACAGAGATGCCCAGGCGAGATTCAGATTTGAAAGCT 2152  
2077 TCCAGGGTTATAGCCAGGACAAAATTTGGGAAGAGATATGCAAGTAAATATGTGACCTAAAGT 2136  
2153 ACCTGCGCTTCAAGCTTCAAGTTTCAGGAGATCCAGAGCTCATCTTAGATCTCAAGT 2212  
2137 ATATCTGTATCAGCAAAACCCACAAATATATGGAAGAGATTAAGAAATGAGATTCCTTTGAAG 2196  
2213 ATGTATTTGATTTAGCAAAACCAACGAGTGTGATGAGCTGAGGCGAGAAATTTCTTACAG 2272  
2197 GTTTTCGATCTTTTTCGAAGATTTCTTCTCTATGCAAGGAAATGGAAGAAATCCGAAGAC 2256  
2273 GGTTCGATGCTCTTCTTGGAAATTTCTGAAGTGCATGCAAGGAAATGGAACCCGATCAGCGTC 2332  
2257 AGTGTGGGCAACACATTTGAAGTGGATCCTGATTTGGAGGCTGCCATTTGCTATACAGATGC 2316  
2333 AGGTGGGACAGCAATTTGAGATGCGACGAGTGGGAAGCAGCTTTCACATGCGAGATGA 2392  
2317 AATTGAAGATATTTTACTCATGTTCCAGAGTGGTGTGCTGTGTGATGAAGAACTCTTTAC 2376  
2393 AGCTGCACACGTCATCTCAATGGTGCAGGACTGGTGTGCTCTGGACGAAAGGTGTTAA 2452  
2377 TTGTGGCTTATAAGAAATGTCACAAAGCTGTGATGAGTGCAGTACCAAGTTTTCATATCTA 2436  
2453 TTGAAGCTTACAGAAATGCTGCTGCTGTGTGACACAGTGTCTATGGCGAATTTACTGATG 2512  
2437 GTAGCAAGACAGT---AGTACAATCGTGTGACATAGTTTGGAAACAAAGTCTTACAGAG 2493  
2513 GTGAACGCCAATCACACTCAGTATTTGTGACACTCGGTGGAACCAATCAGATACTGTG 2572  
2494 TATCTGAGGATCTTTGAAGCATATCATCTGCCACTCTCTAGGACCTTGTGGTCTTCAATG 2553  
2573 TTTCCCAAGAAAAAGTTAGCAATTCACCTCCCAATTTCTCGCTTGTGTCAGGTTTGCATG 2632  
2554 TACGTTTAAAGCAGCTGGTGTCTGTTTCAAGACTGCAATGTTGTGCTTTTGGAGGACT 2613  
2633 TATTGTTAAGCAAAAGTGAAGTGGCATATAAATTTCCAGAGCTCCTACTCTAAGTGAAC 2692  
2614 TTCAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTTGTCTGCTGTTGGTTGCCAGGTTG 2673  
2693 TGAGCCCAACCAATGTTGATAGAACATCCTCTTAGATGCTTGTCTTATGCTCAAGTGC 2752  
2674 TTGCTGAGATGTGGGAAGAAATGGAATGCTCTCTTATTTAGCCAGGTGTTTATTTACCAAG 2733  
2753 ATGCTGGGATGTGGGAAGAAATGGCTTCTCTAGTAAATCAGATCTATTACTACCATA 2812  
2734 ATGTTAAGTCAGAGAGAAATGATGATAAGATATCATCATGCTTCAGATTTGTTGAT 2793  
2813 ATGGAATGAGCGGAGAGATGTTTGAAGAGACATAGTATGCTTCAGACAGTGTCT 2872  
2794 CTTTAAATGGATCCCAATTAAGTTCTTTGTTACTTGGTACTTTTTCAGAGGTATGAACTTGCAGG 2853  
2873 CCATGATGGACCCAAACCACTTCTCTGATGATCATGCTCAGCGCTTTTGAACCTCTATCAGC 2932  
2854 CTTTAAACAGACCATATCTACAAAG-----ACAGGATTTGA 2892  
2933 TCTTCAGCACCCCTGACTATGGGAAGATTCAGTTCTGAGGTTTACCCATGAAGCAGCTCG 2992  
2893 TTAACATATAATACACTAATAGAGAAATGCTTTCAGGTCTCTCATCTATATTTGTTGGTGT 2952  
2993 TTCAGAGACACACTCTGATCGAAGAGATGCTCTACTCATCATCATGCTTGTGAGAG 3052  
2953 AGCGTTATGTACCTGGAGTGGGAAATGTGACCAAGAGAGGTCACAATATGAGAGAAATCA 3012

Db 3053 AAGATTCAAACCTGGGTTGGACAGTGGCTGCCACAGATGAATCAAGAGGAGATTA 3112  
Qy 3013 TTCACTGCTTTGCAATGAACCCATGTCACACAGTGCCATTGCCAAAATTTACTTGAGA 3072  
Db 3113 TCCATCAGTGTAGCATCAAGCTATAGGCTACAGTGAGCTGTGTAAGTCTCTGCTCGAAG 3172  
Qy 3073 ATGAAATTAATCAAACTGGCTTAGAGATGTCATAAACAAGTGGCCACATTTAGAAAC 3132  
Db 3173 ATGAAACAAGAGACCGGCATGGAGAGCGTCAATCGAGTCCGTTGCACATTTCAAGAAAC 3232  
Qy 3133 CAGGTGTATCAGGCCATGAGGTTATGAATCAAAAGATGAATCACTGAAAGACTTCAATA 3192  
Db 3233 CTGGCTCACAGGCGAGGCATGTATGAGTGAAGCCAGAGTGTGCCAAAGATTCAACC 3292  
Qy 3193 TGTACTTTTATCAATTAATCCAAAACCCAGCATAGCAAGGCTGAAATATGCAAGAAATA 3252  
Db 3293 TGTATTTTATCATTTCTCCAGGCGAGAGCAGTCCAAAGCGAGAGGAGTCAAGCGGAAT 3352  
Qy 3253 GGAGAAACAGAAACAAGATGAAGCAATTCGCGCCACCACTCTGTAATTTCTGCC 3312  
Db 3353 TGAAGAAGAGAAATAAAGAGATACAGCACTCCCTCCCGGCTTTGCCACCGTTCTGCC 3412  
Qy 3313 CTGCTTTCAAGCAAGTGAATTAACCTTTCAACTGTGATATCATGATGTACATTTCTCAGA 3372  
Db 3413 CTTTGTTCGGAGTCTGTTAAATCTTTCAGTGTGACGTGATGCTGTACATCATGGGA 3472  
Qy 3373 CGTATTTGAGGGGCAATAGACACAGATTTCTAACTTGTGGACCGAAGGATGTCCAAA 3432  
Db 3473 CGATCTCGAGTGGCTGTAGAGCATCAGGGTCTGCTGCTGAGTCCATGCTACAGA 3532  
Qy 3433 TGGCTTTTATATTCGTCATTTGGGTTTACTAGAAAGAGAGCAAGCTTCAAAAGCTC 3492  
Db 3533 GGGTGTCTGATTTGATCGGGATGGCTCTCCAGGAAGAGAACCACTTGGAGAACGCGG 3592  
Qy 3493 CTGAAGAAGAAGT---AATTTGACTTTTATCATAGGCTTCAAGATTTGGAAGTTGCA 3549  
Db 3593 TGAAGGGCACTGACAGACCTTCACTTCAACAGAGATTTCAAGCTGTGTATGAC 3652  
Qy 3550 -----CCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCAAGTTAG 3600  
Db 3653 CACATACTCCCGAGACTCTAGCTATGCTGGAGACCTTGACAGAACGCCCTCCCTGG 3712  
Qy 3601 AAGGCCAGAGACATGATATAGTGGATATCTCAGATGTTTGAACACAGTGAAGCATTA 3660  
Db 3713 AAGCCACAGAGACATGATCAGGTGGTGTCTAAAGATGTTTAAATGAATTAAGAGATA 3772  
Qy 3661 GAGAAAAATCTGTTTAAATTTAGCAACACATCAGGATCGAATCTATTAAAGATGATG 3720  
Db 3773 GAGAGTTTATCCAGACGCTGTGGCC-----GAGCGAGGGAACCATATGAGAGG 3826  
Qy 3721 AGATTACTCATGATAAAGAAAGCAGAACGAAAGAAAGAAAGCTGTAAGCTGTAGGCTAC 3780  
Db 3827 AGAGCTCAAGAGACAAGCAACAGCAGAGAGAGGAAAGAAAGCCGAGATGCCAGACTGC 3886  
Qy 3781 ATGCGCAGAGATCATGCTGATGCTGCTGCTTTACAGAAATACTTCAATTTGAATCTATA 3840  
Db 3887 GCCGGGAGAAGATCATGCGCCAGAGTGTCTGATGTCAGCGGCATTTCAATGACGAAACA 3946  
Qy 3841 AACTCATGTATGACAAATACATCAGAAATGCTGGGAAGAGATTTCCATTTATGGAGAG 3900  
Db 3947 AAGAGCTCTCCAGCAGACCTTAGAGCTGGACACCTCTGCTCTGCGCACTTTGAC----- 4002  
Qy 3901 AGAGCAACCCAGCAGCTCAGTGAATCTAGAAATTTGGGTCCTTAAACGGGGTCCAT 3960  
Db 4003 ---AGCAGCCCTCCGTTTCAAGCGAGCTCTTACAGCACTGGGCCACGACACAGACAGG 4060  
Qy 3961 CTGTTTCTGAAGAGAGTGTGAGTGTGATCTTTTGGCAAGAAACAGGAGGTGAATA 4020  
Db 4061 TCCCTGAACCCAGACAGTTTGTCACTGTATATATGTTCAAGAGAGGCAAGAGGTGACTG 4120  
Qy 4021 TAGAAATTAATGCCATGTTATTCGGCTGTGTCAGAAATCTACTGCTTAAACCCAGC 4080  
Db 4121 TGGGAAGCAGGCGGATGCTTGGCAGCGTTTGTTCAGAGGTCAACGGTTCTGTCAAAG 4180

Qy 4081 ACAGGGGAAAACCCATAGAACTCTCAGGAGAAAGCCCTAGACCCACTTTTTCATGATCCAG 4140  
Db 4181 ACAGAGCAAAAACCATCGCGACCCAGAAAAATAT---GATCCATTATTTATGCAACCCCG 4237  
Qy 4141 ACTTGGCATATGGAACCTTATACAGGAAGCTGTGGTCATGTAATGACAGCAGTGTCTGCG 4200  
Db 4238 ATCTGTCTTGTGGGACACACACTGSCAGCTGTGGGACGTTATGATGCCCATTTGTTGC 4297  
Qy 4201 AGAAGTATTTTGAAGCTGTAC-----AGCTGAGCTCTCAGCAGCGCATTTCAATGTTG 4251  
Db 4298 AAAGGATTTTGAATTCGGTTCAAGCAAGGAGCAGCAAGGAGCAGCGCTCGCTTGC 4357  
Qy 4252 ACCTT---TTTGAATTTGAAAGTGAATATCTTTGCTCTCTTTGCAAAATCTCTGTGCA 4308  
Db 4358 ACATAGCTACGATGTAGAGAATGCGAGTTCCTGCGCGCTCTGTGAGTGCCTGAGCA 4417  
Qy 4309 ATACTGTGATCCCATTTATCTTTTGAACCTCTTCAAAAGATAAACAAGTGAATGACATG 4368  
Db 4418 ACAGGTTATCCCC-----CTGCTGCTTCTCCAGGAGCATCTCAGCAGGAGTTAA 4471  
Qy 4369 CTCTTGCTCAACTTTTGAACCTGCGCACGCTGGATAACAGTGTCTGCGCCAGAATATCAG 4428  
Db 4472 ATTTTTCAGACCAACAGATCTGSCACAGTGGACGAGAGCAGTAACACAGCAGATAAAGG 4531  
Qy 4429 GTTATATATTAAGACATGCTAAAGGAGAAACCCAAATTCATTTCTTTTAAATCAAGAA 4488  
Db 4532 TGGTCCAGATCTGAGGAGAAAGCAATGCTGTGACACGCTCTTCTTTCAGAGGACACAG 4591  
Qy 4489 TGGGAGATTTCTACTTTTGGAGTTCATTCATCTCAGTGTGTTGGGTTTTCAGTCTTCA 4548  
Db 4592 AAGCCATGAATATTAATACCGATCCCGAAGCTTCAAGGCTTGAATTTTATCTCAGAAC 4651  
Qy 4549 AATATTCAAAATAGCATCAAGGAATGGTTATCTCTTTGGCCACCAACATTTATAGAAATG 4608  
Db 4652 CATATCTGATAGCATAAAGAAATGTTAAACGATTTGGAAACGCTCTTACAGGTGG 4711  
Qy 4609 GATTGAAAGTGCACCTGATGAAAGGATCTCAGATGCCCATGCTGACCTTGAGCACT 4668  
Db 4712 GACTGAAGGTTTCACTTAATGAAGGTGACCCCGGTGTCGCCATCTGCTGCTGGGGGAG 4771  
Qy 4669 GCGCTTTCATCTCCAGCAATTCGAAATCTATTGAGGAGTGAAGGAAACCTCTGTTTG 4728  
Db 4772 GTGCATACACATCAGAGCATAGAAAGAAATTTTGAAGTGAAGGAGGAGGCTGTTTTG 4831  
Qy 4729 GAGCACTTCAAAATAGGAGCATTAATGCTGGAAGCAATTAATGAGTTTTCAGTTGAC 4788  
Db 4832 GACCTCTGCTTGTAGACTGGAACGACTGCTCAGGTGTTAAACAGGTTTTCAGCAGCAC 4891  
Qy 4789 AGAGGATTA CTTGCTCCTCAGGTCTGATPACAGAAACATCTGGTTCGTCTTCTATCAGTTG 4848  
Db 4892 ATTGACAGTGGCGTTACTTCTCTGGTACAAAGGACACTTCTGTAACCTCTTTGATCCT 4951  
Qy 4849 TTCTTCTTAACATAAATCAGAAATACACATGCTCTCTCTATAGATCTGTTTCATG 4908  
Db 4952 TGGTGCCTTAGTGAACAGCTATGAAAGCCTCCCGTGCATACATAGACATCGACATGTTTCACT 5011  
Qy 4909 TTTTGGTGGGTGCTGTGTTAGCAATTCCTTGTATTTGGGATGACCTCTGTTGATCTGC 4968  
Db 5012 TGCTGTTGGGCTGTGCTCGCTTTCAGCTCTGCGAGTCTGCGAGT-----TCAGG 5056  
Qy 4969 AGCCTTCTCAGTTAGTTTCTTCTATAACCACTTTTATCTCTTCCATTTGATCACTATGG 5028  
Db 5057 ATTTTTCAGGAAGCAGCTGCGCCACTGGGACCTGCGACATCTTCCATCTGTTTACATGG 5116  
Qy 5029 CACACATGCTTACATCTACTTACAGTAGACACAGGCTTA---CCCCCTGCTCAGGTTTC 5085  
Db 5117 CACACATGTCACAGATCTTACTTACTCTCAATGACAGAGAGAAATGGCATGGATCAAGAGA 5176  
Qy 5086 AAGAAGCAGTGAAGAGGCTCATTTCCGCTCTCTTTTTCAGGAAATTTCTCAATATA 5145  
Db 5177 ATCCCATGCGGGAAGAGAACTGGCCATTTCTCTTTGCAAAAACATTTCCACCATATA 5236

```
QY 5146 CAAAGTGGCTCCATGGGTGTGATATTCCTGGCTGGTATTTGTGGTCTCACTGAAGAATG 5205
Db 5237 CTGGAAGTGCCTTGAAGAGCCCTCGGCTGGCACTGTGGAGGAGCGTCCGGGCGG 5296
QY 5206 GCATCACCCCTTATCTCGCTGTGTGCTGATTTTTCATATTTTCACTATTTTCTGGGGTAACTC 5265
Db 5297 CCATCATGCTTTCTCTCAAGTGTCTGTCTGTGTTTTCACATATTTTAAATGGAGTCCGGG 5356
QY 5266 GCGCTGAGAACTGATACCAATTTCTGAGAGGAGAGTACAGTGCATCTGTAGCTATC 5325
Db 5357 CCCCTCCAGACCTTCA---AGTTTCTGGAACAAGCCATTTTGAACACTTATGTGTAACACC 5413
QY 5326 TATCTTTTACCTACAAATTTGTCTGCTCTTCCAGGAATATTTGGGATACTGTGAAGCCCT 5385
Db 5414 TTTCTTACCAACCACTCACTTACCTTTTTCAGAAACACAGTGCATCATGAATCTCC 5473
QY 5386 TGCTCCAGAGTGTGTGCAGATCTGCTTACTAACTGTGTGAAGCAAAAAACACCG 5445
Db 5474 TGAATTGAAGTTGGTGCCAGAACAGTGAAGTTAAACGGTATCTAAATGGCGAGAGGAG 5533
QY 5446 TGGTCAGGTACCTAGAAAGAAATAGTTTGTATAGAGCTTCTCATGACTATAGCTGCC 5505
Db 5534 CGATAAGCTTACCCAGAGAGCTAACAACTGTATAGACCTTTCCAGAGGATTTACAGCGCC 5593
QY 5506 TCCTGAATCAAGCTTCTCATTTTCAGTGCACCGTCTGCAGATGATGAGCGAAAGCATC 5565
Db 5594 TATTTAACCAAGCATCCAAATTTCTGTGCCCAATTCAGTGGCGACAGAGCAGAGCTC 5653
QY 5566 CTGCTCTGCTGCTTTTCTGTGGGGCTATATATGTTTCTCAGAACATTTTGTGCCAGGAAA 5625
Db 5654 CTACTCTGTGCTCGTGTGGGAGTCTCTCTGCTCTCAGAGTTTACTGCTGCCAAGCTG 5713
QY 5626 TTGTGAACGGGGAAGAGTGTGAGCTGTGATTTTTCAGCACTTCACTGTGGAGCGGAG 5685
Db 5714 AGCTGGAGGTGAGGACGTCCGAGCTGCACAGCACACACCTACTCTCTGCGGCTCCGGGG 5773
QY 5686 TCTGCATTTTCTTAAATACAGAGATCCGAGTGTCTGCTGCTCAAGCTAAAGCCAGAG 5745
Db 5774 CCGGCACTTCTCTGAGAGTCCGGGAATGTTCAGGTCTATTTTACTGCGCAAAACCAAG 5833
QY 5746 GCTGTGCTTATCCAGCTCTTACTTGGATGATATGGAGAAACAGACCCCTGCGCTGAAGA 5805
Db 5834 GATGTTTTTATCTCTCTCTACCTTGAGACTATGGAGAGACCCAGCAGGACTCAGAC 5893
QY 5806 GGGGCAACCCCTTCATTTATCTGTGAGCGGTATCGAAGCTCCATTTGTCTGGCAAC 5865
Db 5894 GAGGAAATCTTTTACATTTATGCCAAGAGCGGTTTTCGAAGATCCAGAGCTCTGGCAGC 5953
QY 5866 AACACTGCATTTATAGAAGATTTGCTAGGAGCCAGAGACTAATCAGATGTTTATTTGGAT 5925
Db 5954 AGCATAGTATCACAGAGGATTCGGACAGCGCGCAGAGGCTTACCGACCCCTGGTCGGAA 6013
QY 5926 TCAACTGGCAGTTTACTGTGAGCTC 5949
Db 6014 TTGACTGGCAGCATTTATAATCGC 6037
```

## RESULT 13

```
US-10-755-889-634
; Sequence 634, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 634
; LENGTH: 6158
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-634
```

Query Match 15,68; Score 981; DB 19; Length 6158;

Best Local Similarity 55.98; Pred. No. 3.8e-259; Indels 90; Gaps 12;  
Matches 2157; Conservative 0; Mismatches 1610;

```
QY 2130 CTAAAGTATATCTCGATCAGCAAAACCCACAAATATGGACAGAAAGATTAAAGAAATCGATTC 2189
Db 1 CTCAAGTATGTGTTAAATTAGCAAAACCAACTGAATGGTCAGATGAGTCGAGCAGAAGTTC 60
QY 2190 CTTGAAGGTTTTCATCTTTTGAAGATCTTTTACCTGTATGCGAGGATGGAAGAATC 2249
Db 61 CTAGAAGGGTTTGTATGCTTTTGGAAATTAATAAATGTATGCGAGGGAATGGATCCAAAT 120
QY 2250 CGAAGACAGGTTGGGCAACACATTTGAAGTGCATCTGATTGGGAGGCTGCCATTTGCTATA 2309
Db 121 ACAGCTCAAGTAGGACAAACATATTGAATGGAAACAGAGTGGGAAGCAGCTTTCACACTA 180
QY 2310 CAGATGCAATTTGAAGAATATTTTACTCATGTTTCAAGAGTGGTGTGTGTGATGAAGAA 2369
Db 181 CAATGAAATTTAACACATGTCATTTCAATGATGCAGGACTGGTGTGCTTCAGATGAAAAA 240
QY 2370 CTCTTACTTGGGCTTATAAAGAATGTCAAAAGCTGTGATGAGTGCAGTACCAGTTTC 2429
Db 241 GTGTTAATCGAAGCTTCAAGAAATGTCTGCTGTACTGATGCGAGTGTGCATGGTGGTTAT 300
QY 2430 ATATCTAGTAGCAAGACAGT---AGTACAATCGTGTGACATAGTTTGGAAACAAAGTCC 2486
Db 301 ACTGATGTGAAACAGCCCAATCACACTAAGCATTTGTGACATTCAGTGGAAACTATCAGA 360
QY 2487 TACAGAGTATCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGTGGT 2546
Db 361 TACTGTGTTTCCCAAGAAAAAGTTAGCATTCACCTCCAGTTTCTCGCTTACTTGCAGGT 420
QY 2547 CTTCAATGACGTTTAAAGCAGGCTGGTGTCTTTTCAAGACTGCATGAATTTGTGTCTTTT 2606
Db 421 TTACATGTTATTTAAAGCAAAAGTGAAGTGGCATATAAAATTTCCAGAGCTCTTACCTCTA 480
QY 2607 GAGGACTTTCAAGTAGAGGTACTAGTGAATATCTTTTACGTGTGTGTGTGTGTGTGTC 2666
Db 481 AGTGAACCTTACCCACCACTGTTGATGAAGAACCCCTCTTAGATGTCTTGTCTGTGTGTC 540
QY 2667 CAGGTTGTGTGAGATGTGCGAAGAAATGGACTGTCTTATTTAGCCAGGTGTTTAT 2726
Db 541 CAAGTACATGCGGAAATGTGGAGAAGAAATGGGTTCTCTCTAGTAAACAGATTTATTATAC 600
QY 2727 TACCAAGATGTTAAGTGCAGAGAAGAAATGATGATAAGATATCATGCTTTCAGAT 2786
Db 601 TACCATAATGTGAATTCAGACGTCGAGATGTTTGAACAGGATGTAGTAATGCTTTCAGACA 660
QY 2787 GGTGCATCTTTAATGGATCCCAATAAGTTCTGTGTACTTGTGTACTTCAGAGGTATGAACCT 2846
Db 661 GGTGTCTCCATGATGATCCAAATCTTTCTGTGATGATCATGCTCAGCGGCTTTGAACCT 720
QY 2847 GCCGAGGCTTTTAAACAGACCATATCTACAAAA-----GACAG 2885
Db 721 TATCAGATTTTTCAGTACTCCAGACTATGGAAGAAAGATTTAGTCTGTGAGATTAACCATAG 780
QY 2886 GATTGTATTAACAAATATAATACACTAATAGAAGAAATGCTTTCAGGCTCTCATCTATATT 2945
Db 781 GATGTTGTTTACGACAGAACAAATCTCTAATAGAAGAAATGCTTATACCTCATTTATATGCTT 840
QY 2946 GTGGTGAAGCTTATGTACTCTGGAGTGGGAAATGTGACCAAGAAAGAGGTCACAATGAGA 3005
Db 841 GTTGAGAGAGATTTAGTCTCTGGAGTTGACAGAGTAATGCTACAGATGAATCAAGCGA 900
QY 3006 GAAATCATTTCTTGTCTTTGCAATTTGAACCCCAATGCCACACAGTGCATTTGCCAAAAATTTA 3065
```

Db 901 GAGATTATCCATCAGTTGAGTATCAAGCCTATGGCTCATAGTGAATTGGTAAAGTCTTTA 960  
Qy 3066 CTTGAGAAATGAATAAATGAAATCTGGCTTATGAGAAATGTTCATAAAACAAAGTGGCCACATTT 3125  
Db 961 CCTGAAGATGAGAAACAAAGGAGACTGGCATGGAGAGTGAATCGAAGCAGTTGCCATTTTC 1020  
Qy 3126 AAGAAACCAAGGTGATCAGGCCATCGAGTTTATGAACCTAAAGATGAATCACTGAAAGAC 3185  
Db 1021 AAGAAACCTGGATTAAACGAGCAGAGCATGTATGAACCTGAACCCAGAAATGTGCCAAGAG 1080  
Qy 3186 TTCAATATGTATCTTTTATCATTTACTTCCAAAACCCAGCATAGCAAGCTGCAACATATGCAG 3245  
Db 1081 TTCAACTGTGATTTCTATCACTTTTCAAGGCGAGAACAGTCCAGGCGAAGAGCGCAA 1140  
Qy 3246 AAGAAAGGAGAAAACAAGAAAACAAGATGAAGATGCGGCCACCACTCCTCTGAA 3305  
Db 1141 CGGAAATGAAAGACAAATAGAGAAAGATACAGCACTCCACCTCCGGTGTGCTCCA 1200  
Qy 3306 TTCTGCCCTGCTTTCAGCAAGCTGATTAACCTTCTCACTGTGATATCATGATGTACATT 3365  
Db 1201 TTCTGCCCTGCTTTGCAAGCCTGGTTAAACATTTTGCAGTCAGATGTCAATGTGTGTCATC 1260  
Qy 3366 CTCAGACCGGTATTTGAGCGGCAATAGACACAGATTTCACTTTGTGACCGAAGGGATG 3425  
Db 1261 ATGGGACAAATTCGCAATGGGCTGTGGAAACATAATGGATATGCTTGGTCAGATGCCATG 1320  
Qy 3426 CTCCTAAATGGCTTTTCATATCTGGCATTTGGGTTTACTAGAGAGAAAGCAACAGCTTCAA 3485  
Db 1321 CTGCAAGGGGTGTACATTTAATTTGGCATGGCACTACAAAGAAAGAAAACAACATTTAGAG 1380  
Qy 3486 AAGCTCCTGAAG----AAGAACTAACATTTGACTTTTATCATAGGCTTCAAGNTTGGGA 3542  
Db 1381 AATGTACGGAAAGAGCATGTAGTACATTTTACCTTCACTCAGAGATATCAAAACCTGGT 1440  
Qy 3543 AGTTCAGCCATCAAT-----ATACAAATGCTTTTGGAAAACTCAAAGGAATTCOC 3593  
Db 1441 GAAGCGCCAAAATTTCTCCTAGCATATAGCTATGCTGGAAAACTACAAATGCTTCC 1500  
Qy 3594 CAGTTAGAGGCCAGAGGACATGATAAGCTGGATATCTTCAGATGTTTGAACAGTGAAG 3653  
Db 1501 TACCTAGAGTCCACAAGACATGATTCGGTGGATATTGAAGACTTTTAAATGCTGTAAA 1560  
Qy 3654 CGATTAGAGAAAATCTTTGTTAATTTGTAGAACCAATCAGGATCGGAATCTATTAAG 3713  
Db 1561 AAGATGAGGGAGA-----GTTCACTACAGTCCCGTGGCAGACAGAGAGAAACCATTA 1614  
Qy 3714 AATGATGAGATTACTCATGATAAGAAAAGCAGACGAAAGAAAAGCTGAAGCTGT 3773  
Db 1615 ATGGAGAGAGTTCAAGGGAACAAAGACAAAGCTGAGAGAGAGAAAAGCAGAGATTGCC 1674  
Qy 3774 AGGCTACATCGCAGAGATCATGGCTCAGATGTCTGCGCTTACAGAAAACCTTCAATTGAA 3833  
Db 1675 AGACTGCCAGAGAAAAGATCATGGCTCAGATGTCTGAAATGCAAGGGCATTTTATTGAT 1734  
Qy 3834 ACTCATAACTCATGTATGACAAATCATCAGAAATGCTCGGAAAGAGAAATTCATTAAG 3893  
Db 1735 GAAACAAAGAACTCTTTTCAGCAGACATTAAGAACTGGATGCTCAACCTCTGCTGTCTT 1794  
Qy 3894 GAGGAAGAGAGCACCACAGCAGTCACTACTCTAGAAATGCTTTGGTCTTAAACGG 3953  
Db 1795 GATCA-----TAGCCCTGTGGCTTTCAGATATGACATTTACAGCACTGGGTCCTCAACAA 1848  
Qy 3954 GGTCCATCTGTATCTGAAAGAGGTGTGAGCTGCATCTTTGGCAAGAAAGACAGGAG 4013  
Db 1849 ACTCAGGTTCTGAAACAAAGCAATTCGTTACATGTATATTGTGTCAAGAGGAGCAGAA 1908  
Qy 4014 GTGAAAATAGAAAATAATGCCATGTATATCGGCTGTGTCCAGAAATCTACTGCCCTTA 4073  
Db 1909 GTTAAAGTGAAGAGCAGGCAATGCTTTGGCAGCAATTTGTTTCAGAGATCAACTGTATTA 1968  
Qy 4074 ACCCAGCAGGGGAAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTCATG 4133  
Db 1969 TCAAAAAACAGAAAGTAAATTTTATTCAGAGATCCAGAAAAATAT---GATCCCATTTTCATG 2025

Qy 4134 GATCAGACTTGGCATATATGAACTTATACAGAGAGCTGTGTCTATGTATATGACGCGAGTG 4193  
Db 2026 CACCCTGATGTCTTGTGTGAAACACACACTAGTAGCTGTGGGCACATTTATGATGCCCAT 2085  
Qy 4194 TGCTGGCAGAAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTGTAC 4253  
Db 2086 TGTGGCAAGGTATTTTGTATTCCTTCAAGCTAAAGAACAGCGAAGGCAACAGATTA 2145  
Qy 4254 CTTTT-----TGACTTGGAAAGTGGAGATATCTTTGCCCTCTTTTGCAAATCT 4301  
Db 2146 CGCTTACATACGAGCTATGATGTAGAAAACGAGAAATTCCTTTGCCCTTTTGTGAATGC 2205  
Qy 4302 CTGTGCAATATCTGTGATCCCATTTATCTTTTGGAACTCTCAAGAAATTAACAGTGAGAT 4361  
Db 2206 TTGAGTAATATCTGTTATTTCC-----TCTGCTGCTTCTCTCAAGAAATATTTTAAACAAC 2259  
Qy 4362 GCAGATGCTCTTGTCTCAACTTTTGAACCTGGCAGCGTGGATACAGACTGTTCTGGCCAGA 4421  
Db 2260 AGGTTAAATTTTTCAGACCAACCAATCTGACTCAGTGGATTAGAACAAATATCTCAGCA 2319  
Qy 4422 ATATCAGGTTATTAATATAGACATCTTAAAGGAGAAAAACCCAAATTCCTATTTTCTTAAAT 4481  
Db 2320 ATAAAGCATTTACAGTTCTTAGGAAAGAGAAAGTACTCTTAATAATGCTCTACAAAG 2379  
Qy 4482 CAAGGAATGGGAGATTTCTACTTTGGAGTTCCATTCCTCTGAGTTTTTGGCGTTGAGTCT 4541  
Db 2380 AATTTCAGAAAATGTGATGAATTAACAGCTCCCTGAAGGGTTTCAGGCTCTGATTTTGGTCT 2439  
Qy 4542 TCGATTAAATATCAAAATAGCATCAAGAAATGTTTATCTCTTTGGCCACCAACATTTAT 4601  
Db 2440 AAGATCCCTTATTTTCGAGAGCAATAAAGAAATGCTAAACGCAATTTGGAACTCTCCTAC 2499  
Qy 4602 AGAATTCGATGAAAGTCCCACTCATGAAAGGGATCTCTCGAGTCCCATCTGTCAGCTGG 4661  
Db 2500 AAGTGGGACTAAGGTTTCATCCCAATGAGAGGATCTCTGTTTCCCATTAATGTTGG 2559  
Qy 4662 AGCCTCGCTTTCATATCCAGCAATTTGAAATCTATTGGGAGATGAAGAAACCT 4721  
Db 2560 GGTAGCTCGGCTACACCATCCAAAGCATAGAAAGAAATTTTGAAGTGAAGATAAACCA 2619  
Qy 4722 CTGTTTGGAGCACTTCAAAATAGGCGCATTAATGCTCTGAAAGCAATTAATGCAATTTGCA 4781  
Db 2620 TTGTTTGGTCTCTTACCTTTCAGAGCTGATGATCTGTCTAGTCAATTCAGAGATTTGCC 2679  
Qy 4782 GTTGCAAGAGATTTACCTGCTCTCAGGTCCTGATACAGAAACATCTGTTCTGTTCTTA 4841  
Db 2680 GCAGCACTGCAAGCTGGCATCAGTTTCAGTGGTGGAGGACATTTTGTAAACCTTTT 2739  
Qy 4842 TCAGTTGTTTCTTCAACATAAATTCAGAAATACACCATGCTTCTGTCTATAGATCTG 4901  
Db 2740 GCATCACTGGTGCCTTAATGACAGCCCATGAGGAACTTCCATGCATATTAGATATTGACATG 2799  
Qy 4902 TTTCAATGTTTGGTGGGCTGTGTGATGATTCATTCCTGATCCCATCTGTTATGCGATGCCCTGTT 4961  
Db 2800 TTTCAATTTATTTGGTGGGCTTGGTGTGCTTTCATTTCTGCGTTCGCACTG----- 2846  
Qy 4962 GATTCGAGCTTCTTTCAGTTAGTTCTTCTTATTAACCACTTTATCTCTTCCATTTGATC 5021  
Db 2847 ---TCAGGATTTTTCAGGATCAGCCTTGGCACTGGAGACCTTCAATTTTCAATCTGTT 2904  
Qy 5022 ACCATGGCACTATGCTTTCAGATATCTTTCAGTAGACAGAGCC---TACCCCTTGTCT 5078  
Db 2905 ACTATGGCACAACATCATACAGATCTTACTTACCTCATGTACAGAGAGAAAGGCGATGAT 2964  
Qy 5079 CAGTTTCAAGAGACAGTGAAGAGCTCATTTCCGATCTTCTTTCTTTTCAGAGAAATTTCT 5138  
Db 2965 CAAGAAAATCCCTCTTGTGAAGAAAGAAATCAGCAGTTCTTGTCTTTGTATTAACACTTCAAC 3024  
Qy 5139 CAATATCAAGTGGCTCCATTTGGGTGTGATATTTCTGCTGTGATTTTGTGGGTCTCACATG 5198  
Db 3025 CAGTATACGGGAAGTGCCTTGAAGAAATACCATCCGCTGGCATCTGTGGAGGAGTGTCT 3084



Qy	5199	AAGAATGGCAATCA	CCCTTATCTTCGCTGCTGCA	TGTTTTTCCACTATTTACTTGGG	52558
Db	3085	AGAGCTGGAATCAT	GCCTTTCTCCGAAAGTGTTCTGCTTATTTTTCATTACTTAAATGGA	3144	
Qy	5259	GTAAC	TCGGCTGAGGAAC	TGCATACCAATTCCTGCAGAGGAGTACAGTGCAC	5318
Db	3145	GTTCTCTCCCA	CCCGACATTC	---AGTTCCTGGAACAAGCCCA	3201
Qy	5319	AGCTATCTATCT	TTTACCTACA	CAATTTCTGCTCTTCCAGGAATATTCGGGATAC	5378
Db	3202	AGCTATCTTTCC	CTTACCAACCACTCAT	TTGGCTTTTCAAGGAATATGAGATAATG	3261
Qy	5379	AGGCC	TTGTCTCAGAGGGTGTGCAGAT	CTCTGCCCTTATAA	5438
Db	3262	AA	TTCACTGATGAA	AGTTGGTGGCGGTAA	3321
Qy	5439	AACAC	CGTGGTCAGGTAC	CTTAGAAGAAATAGTTTGATAGAGCTTCCTGATGACTAT	5498
Db	3322	AGAGATGCTAT	AAGATATCCAAGAGAA	ATCTAACAAATTAATAAACCTTCCAGAGGATTAC	3381
Qy	5499	AGCTGCCTCT	CGTAATCAAGCTTCTCAT	TTTCAGGTGCCCA	5558
Db	3382	AGCAGCTCAT	TATCAAGCAT	CCAAATTTCTCGTGCCCGAATTCAGGTGGTGATAGAGC	3441
Qy	5559	AAGCAT	CTGTCTCTGCTTTTCTGTGGGGCTA	TAATATGTTCTCAGAACATTTGCTGC	5618
Db	3442	AGAGCCCA	CAACTGTGTGCTTGTGTGGGAT	CTCTGTGTCTCCAGAGTTACTGTCTGC	3501
Qy	5619	CAGGAAT	TGTGAACGGGGAGAGGT	TGGAGCTTGCAATTTTTCACGCACTTCACTGTGGA	5678
Db	3502	CAGACTGAA	CTGGAGGGGAGGAT	TAGAGCCTGCACAGCTCACCTACTCTCTGTGGC	3561
Qy	5679	GCCGGAGTCT	GTGCAATTTTCTTAA	AAATCAGAGAATCCGAGTGGTCTCGTTGTAAGGTAAA	5738
Db	3562	TC	TGGAGTGGGCATCTTCTCAGAGTACGGGA	ATGTCAAGTGTATTTTAGCTGGCACA	3621
Qy	5739	GCCAGAGGCTGT	GCCTATCCAGCTCTTATCTTGA	TGAATGAATATGAGAGAAACAGACCTGTGGC	5798
Db	3622	ACCA	AAAGGCTGTTTTTATCTCTCTCTTACCTT	GATGACTATGGGAGACCGAC	3681
Qy	5799	CTG	AGAGGGGCAACCCCTTCA	TTTATCTCGTGAGGGTATCGGAGCTCCATTTGGTC	5858
Db	3682	CTC	AGAGGGGAAATCTTTTACATTTAT	GCAAGAGCGATTCAGAGAGCTC	3741
Qy	5859	TGG	CAACAACTGCA	TTATAGAGAGATTTGCTAGGAGCCAAAGAGACTAATCAGATGTTA	5918
Db	3742	TGG	CAACCAACACAGTGT	CACAGAGGAATTTGGACATGCACAGGAAGCCATCAGACACTG	3801
Qy	5919	TTTGG	ATTCAACTGGCA	5935	
Db	3802	GTT	CGCATTTGACTGGCA	3818	

RESULT 14  
US-09-529-063-56  
; Sequence 56, Application US/09529063  
; Patent No. US20020102542A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAIKICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; TITLE OF INVENTION: THE BOTH  
; FILE REFERENCE: Q58769  
; CURRENT APPLICATION NUMBER: US/09/529,063  
; CURRENT FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; PRIOR FILING DATE: 1997-10-07  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56

```

; LENGTH: 3327
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-529-063-56

Query Match      12.9%; Score 813.6; DB 9; Length 3327;
Best Local Similarity 55.4%; Pred. No. 5e-213;
Matches 1856; Conservative 0; Mismatches 1409; Indels 87; Gaps 11;

Qy      2632  TGGAAATACCTCTTACAGTTGTCTGGTGTGTGTCGCCAGGTGTGTGCTGAGATGTGGCGAA 2691
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      2692  GAAATGACGTGCTCTTATTTAGCCAGGTGTTTTATTACCAAGATGTTTAACTGTCGAGAGAAG 2751
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      68   GAAATGGTGTCTCTCTAGTAAACCCAGATTTATTACTACCAATAATGTGAAATGCGACGCTG 127
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      2752  AATGTATGATAAAGATATCATCATGCTTCAGATTTGGTGCATCTTTTAATCGATCCCAATA 2811
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      128  AGATGTTTGACAGGATGTAGTAATGCTTCAGACAGGTGTCTCATGATGATCCAAATC 187
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      2812  AGTTCCTTGTACTGGTACTTCAGAGGTATGAACCTTGCAGGCTTTTAAACAAGCCATAT 2871
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      188  ATTTCTGTATGATCATGCTCAGCGCTTTGAACTTTATCAGATTTTCAGTACTCCAGACT 247
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      2872  CTACAAA-----GACCAGGATTTGATTAACAATATAATACAC 2910
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      248  ATGGAAAAAGATTTAGTTCTGAGATTACCCATAAGGATGTTGTTTCAGCAGAACATACTC 307
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      2911  TAATAGAGAAATGCTTCAGTCTCTCATCTATATTGTGGGTGAGCGTATGTATACCTGGAG 2970
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      308  TAATAGAAAGAAATGCTATACCTCATTTAATGCTTGTGTGAGAGAGATTTAGTCTCGAG 367
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      2971  TGGAAATGTGACCAAGAAGAGGTCAATAGAGAGAAATCATTTCACTTGCTTTGCAATG 3030
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      368  TTGCACAGGTAAATGCTACAGATGMAATCAAGCGAGAGATTATCCATCAGTCAGTATCA 427
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3031  AACCATGCCACACAGTGCCATTTGCCAAAAATTTACTCGAGATGAAAAATTAATGAACCTG 3090
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      428  AGCCTATGGCTCATAGTGAATTTGTAAGTCTTTTACCTGAAGATGAGAACAGAGAGACTG 487
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3091  GCTTATAGAAATGTCATAAAACAAATGGCCACATTTAAGAAACCAAGGTGTATCAGGCCATG 3150
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      488  GCATGGAGAGTGAATCGAAGAGCTTGCCCATTTTCAGAAACCTGGATTAACAGGACGAG 547
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3151  GAGTTTATGAACATAAAAGATGAAATCACTGAAAGACTTCAAATATGTACTTTTATCATTTACT 3210
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      548  GCATGTATGAATGAAACCAAGATGTGCCAAAGAGTTTCAACTTGTAATTTCTATCACTTTT 607
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3211  CCAAAACCCAGCATAGCAAGCTGMAATATGAGAGAAAGGAGAAACACAGAAACA 3270
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      608  CAAGGGCAGAACACTCCAAAGCGAAGAACGCGCAACGGAATTTGAAAGAACAATAAGAG 667
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3271  AAGATGAAGCAATTCGCCACCACCACTCTCTGAATTTGCGCCCTGCTTTTCAGCAAAAGTGA 3330
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      668  AAGATAAGCACTCCCACTCCGGTGTGCTCCATTTCTGCGCTCTGTTTTCAGACCTTGG 727
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3331  TTAACTTCTCAACTGTGATATCATGATGATCAATTTTCAGGACGGTATTTGAGCGGGCAA 3390
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      728  TTAACATTTTTCAGTTCAGATGTATGTTGTGTCATCATGGGAAACAAATTTCTGCAATTTGG 787
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3391  TAGACACAGATTTAACTTTGGACCGAAGGATGCTCCAAATGGCTTTTCAATATTTCTGG 3450
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      788  TGGAAACATAATGGATATGCTTGGTCAGATTCATGCTGCAAGGGTGTTAACATTTAATTG 847
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3451  CATTTGGGTTTACTAGAAGAGAAGCAACAGCTTCAAAAAGCTCTCTGAAAG---AAGAAGTAA 3507
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      848  GCATGGCACTACAGAAGAAACAAACATTTAGAGATGTACAGGAAGAGCATGTAGTAA 907
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      3508  CATTTGACTTTTATCATTAAGCTTCAAGATTTGGAGTTTCAGCATGAAT-----A 3558
Db      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
Qy      908  CATTTACCTTCACCTCAGAAGATATCAAAACCTGTGTGAAGCGCCAAAAAATTTCTCTAGCA 967

```

3559 TACAAATGCTTTTGGAAAACTCAAGGAATTCCTCCAGTTAGAGGCCAGAGGACATCA 3618  
968 TACTAGCTATGCTGGAAAAACATACAAAATGCTCCCTACCTAGAGTCCCAAGACATGA 1027  
3619 TAACGTGGATATCTTCAGATGTTTGAACAAGTGAAGCGATTAAGAGAAAAATCTTGTTAA 3678  
1028 TTCCGTGGATATTGAAGACTTTTAATGCTGTTAAAGATGAGGAGA-----GTTTAC 1081  
3679 TTGTAGCAACCATCAGATCGGAATCTATTAAAGATGATGAGATTAATCTCATGATAAAG 3738  
1082 CTACCAGTCCCGTGGCAGAGACAGAGGAACCAATAATGGAAGAGAGTTCAAGGGACAAAG 1141  
3739 AAAAGCCAGAAAGAAAGAAAGCTGAGCTGCTAGGCTATACCTGCCAGAGAGATCAGG 3798  
1142 ACAAGCTGAGAGGAAGAAAGACGAGATTGCCAGACTGCCGAGAGAAAGATCATGG 1201  
3799 CTCAGATGCTGCCCTTACAGAAAACTTCATTTGAAACTCATAACTCATGATGACAATA 3858  
1202 CTCAGATGCTGMAATGCAAGCGCATTTTATGATGMAAACAAAGAACTTTTCAGCAGA 1261  
3859 CATCAGAAATGCTGGGAAAGAGATTCATTTATGAGGAAGAGAGACCCCGCAGCTCA 3918  
1262 CATTAGAACTGATGCTCAACTCTGCTGTTCTTGATCA-----TAGCCCTGTGGCTT 1315  
3919 GTGACTACTCTAGATTTGCTTTGGGTCTTAAGCGGGTCCATCTGTACTGMAAGGAGG 3978  
1316 CAGATATGACACTTACAGCACTGGGCCCGCACAACCTCAGGTTCTTGAACAAAGACAAT 1375  
3979 TGCTCAGCTGCATCTTTGCCAAGAGAAACAGAGGTGAATAGAAATATGCAATGCAATG 4038  
1376 TCGTTACATGATATTGTTCAAGAGGACAGAGATTAAAGTGGAAAGCGGCAATGG 1435  
4039 TATTATCGGCTGTGTCAGAAATCTACTGCCCTTAACCCAGACAGGGGAAAAACCCATAG 4098  
1436 TCTTGGCAGCAATTTGTCAGAGATCAACTGTATTATCAAAAACAGAGTAATTTATTC 1495  
4099 AACTCTCAGGAGAGCCCTAGACCACTTTTCATGATCCAGACTTGGCATATGGAATTT 4158  
1496 AAGATCCAGAAAAATAT---GATCCATTTATTCATGACCCCTGATCTGTCTGTGGAACAC 1552  
4159 ATACAGGAAGCTGTGCTATGTAAGTCAAGCAGCTGTGCTGGCAGAGATTTTGAAGCTG 4218  
1553 ACACGTAGTGTGGGACATATGATGCCCATTTGTGGCAAGGATTTTGTATTCG 1612  
4219 TACAGCTGAGCTCTCAGCAGCGCAATTCATGTTGACCTTTT-----TGACTTGG 4266  
1613 TTCAAGCTAAAGAACAGCAGGAGCAACAGAGATTACGCTTACATACGAGCTATGATGAG 1672  
4267 AAAGTGGAGATATCTTTGCCCTCTTTGGCAATCTCTGTGCAATCTGTGATFCCCATTA 4326  
1673 AAAACGAGAAATTCCTTTTGGCCCTTTTGTGAATGCTTGAATTAATCTGTTATTCCTG 1732  
4327 TTCCCTTGCACCTCAAGATAACAGTGAAGATGCAGATGCTCTTCTCACTTTTGA 4386  
1733 TGCCTT-----CTCAAGAAATATTTTAAACACAGGTTAAATTTTCAAGCAACCA 1786  
4387 CCTTGGCAAGTGGATACAGACTGTTCTGGCCAGAAATATCAGGTTATTAATAAGACATG 4446  
1787 ATCTGACTGAGGATTAAGCAATATCTCAGCAATTAAGCATTAACATGTTCTTAGGA 1846  
4447 CTAAGGAGAAAAACCAATCTCTTTTCAATCAAGGAATGGAGATCTACTTTGG 4506  
1847 AAGAAAGATGACTCTTAATAATGCTCTACAAAGAAATTCAGAAAAATGAGATGAATTA 1906  
4507 AGTTCCATTCATCTGAGTTTGGGTTGAGTCTTCGATTAATTAATTAATCAATAGCATCA 4566  
1907 AGCTCCCTGAAGGGTTCAGGCTGATTTTTCGCTTAAGATCCCTTATTTCTGAGACATAA 1966  
4567 AGGAATGTTATCTCTTTGGCAACAAATTTATAGAAATTTGAAGTGGCCACCTG 4626  
1967 AAGAAATGCTAACGACATTTGGAATCTGTAACCTACAGGTGGGACTAAAGGTTTATCCA 2026  
4627 ATGAAAGGATCTCTCAGTCCCATGCTGACCTGGAGCACCTGGGCTTTCTCATATCCAGG 4686

2027 ATGAGAGGATCCTCGTGTCCATAATGTTGGGGTAGCTGCGGTACACATCCAAA 2086  
4687 CAATTTGAAAAATCTATTGGGAGATGAAGAAAACTCTGTTTGGAGCACTTTCAAAAATAGGC 4746  
2087 GCATAGAAAGAAATTTTGAAGTGAAGATAAACCAATGTTTGGTCTCTTACCTTGCAGAC 2146  
4747 AGCAATATGTTCTGAAGCAATTAATGCAAGTTTGGAGTTGCAAGAGGATTAATCTGCTC 4806  
2147 TGGATGACTGTTTAGGTCAITGACGAGATTTTGGCGCAGCACTGGAAGTGGCATCAG 2206  
4807 AGGTCTCTGATACAGAAACATCTGTTGCTCTTCTATCAGTTGTTCTTCTTAACATAAAT 4866  
2207 TTTGAGTGGTGAAGACATTTTGTAACTTTTTCATCTGTTGCTTAATGACAGCC 2266  
4867 CAGAAGATACACATGCTCTGTCTATAGATCTGTTTCATGTTTGGTGGGCTGTGCT 4926  
2267 ATGAGAACTTCCATGATATAGATATGATGATGTTTCAATTTATTTGGTGGCTTGGTGC 2326  
4927 TAGCAATCCCATCTTGTATTTGGGATGACCTGTTGATCTGCGAGCTTCTTTCAGTTAGTT 4986  
2327 TTGCAATTTCTGCTGCTGCTGCTG-----TCAGGATTTTTCAGGGATCAGCC 2371  
4987 CTTCCTATAACCACTTTATCTTTCATTTGATCACTATGCAATGCAATGCTTTCAGATAC 5046  
2372 TTGGCACTGGAGACCTTCACTTTTTCATCTGTTTACTATGGCAACATATACAGATCT 2431  
5047 TACTTTACAGTAGACAC---AGGCCCTACCTTGTCTCAGGTTCAAGAAAGACAGTGAAGAG 5103  
2432 TACTTTACCTCATGATCAGAAAGATGCGATCAAGAAATCCCTCTTGTGAAGAG 2491  
5104 CTCAATCCGCACTTTCTTTTTCAGAAATTTCTCAATATATACAAAGTGGCTCCATTTGGT 5163  
2492 AATCAGCAGTTCTTGTCTTTGATATAAACACTTACAGTATACGGGAAGTGGCTTGAAG 2551  
5164 GTGATTTCTGCTGGCTGATTTTGGGTCTCACTGAAGAAATGGGATCACTCCCTTATCTTC 5223  
2552 AAATACCATCTGGCTGGCATCTGTGGAGAGTGTGAGAGTGAATATGCTTCTCTGA 2611  
5224 GCTGCTGCTGATTTGTTTCACTATTTTCACTTTGGGTAACTCCGCTCAGGAATCGCAT 5283  
2612 AGTGTCTGCTTTATTTTCACTTTAATGAGTTCTTCCCAACCCGACATCTCAG 2671  
5284 CCAATTTGCAAGAGGAGATGACAGTGCATCTGTAGTATCTATCTTTTACCTACAAAT 5343  
2672 T---TCCTGGAAACAGCCATTTTGAACATTTATGATGCTATCTTCCCTACCAACACC 2728  
5344 TGTTCCTGCTCTTCCAGGAATTTGGGATATGTTAAGGCCCTTGTCTCAGAGGTGGTGT 5403  
2729 TCATTTGCTTTTTCAGAAATAGTGAGATAATGAATTTCACTGATTTGAAGTTGGTGCC 2788  
5404 CAGATCTGCTTACTTAACTGTTTGAAGCAAAAAACACCGTGTGAGTACCTTAGAA 5463  
2789 GTAACAGTGAATTAAGATATCTAGAAAGTGAAGAGATGCTATAGATATCCAGAG 2848  
5464 AAAGAAATAGTTTGTATAGAGCTTCTGATGATCTAGTGTCTCTGATCAAGCTTCTC 5523  
2849 AATCTAAACAATTAATAACCTTCCAGAGGATTAACAGAGCTCATTAATCAAGCATCCA 2908  
5524 ATTTCAGGTGCCAGGCTGCAAGATGATGAGGAAAGCATCTGCTCTCTGCTTCTTCT 5583  
2909 ATTTCTGCTGCCGAAATCAGGTGTGATAAGAGCAGAGCCCACTCTGTGCTTGTGT 2968  
5584 GTGGGCTTACTATGTTCTCAGAACATTTGCTGTCAGAAATTTGTGAACCGGGAAGAG 5643  
2969 GCGGATCTCTGCTGCTCCAGAGATTAATGCTGTCAGACTGAACTGGAAGGGAGGATG 3028  
5644 TTGAGCTTGAATTTTTCACGCACTTCACTGTGAGCGGAGTCTGCAATTTTCTTAAAAA 5703  
3029 TAGGAGCTGCAAGCTCACACCTTCTCTGCTGCTGAGTGGGATCTTCTCTGAGAG 3088  
5704 TCAGAGATGCGGAGTGTCTGTTGAAGGTAAGCCAGAGGCTGTGCTTATTCAGCTC 5763



QY 4099 AACTCTCAGAGAGCCCTAGACCACCTTTTCATGGATCCAGACTTGGCATATGGAACTT 4158  
DB 1496 AGATCCAGAAATAAT---GATCCATTAATTCAGCACCTGATCTGCTTGGGAACAC 1552  
QY 4159 ATACAGAAAGCTGTGGTCAATGTAATCAGCAGCAGTGTGCGCAGAAAGTATTTTGAAGCTG 4218  
DB 1553 ACACCTAGTAGCTGTGGGCACATTAATGATGCCCCATTTGTTGGCAAGGATTTTGTATCCG 1612  
QY 4219 TACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTTTT-----TGACTGG 4266  
DB 1613 TTCAAGCTTAAGAACACGCGAAGGCAACAGAGATTACGCTTACATACGAGCTATGATGTAG 1672  
QY 4267 AAGTGGAGAAATATCTTGGCCCTCTTTGCAATCTCTGTGCAATACCTGTGATCCCATTA 4326  
DB 1673 AAAACGAGAAATCTTGGCCCTTTGGAATGCTTGAGTAACTGTTATCTCTCTGC 1732  
QY 4327 TTCTTTGCAACCTCAAAAGATAAACACAGTGAGAAATGCAGATGCTCTGTGCTCAACTTTTGA 4386  
DB 1733 TGCTTT-----CTCCAGAAATAATTTTAAACAACAGGTTAAATTTTCAGACCAACCAA 1786  
QY 4387 CCTTGGCAGGTGGATACAGACTGTTCTGGCCGAGAAATATCAGGTTAATAATATAAGACATG 4446  
DB 1787 ATCTGACTCAGTGGATTAGAACAAATATCTCAGCAATAAAGCATTAACAGTTCTTTAGGA 1846  
QY 4447 CTAAAGGAGAAACCAATCTCTATTTCTTTAACTCAAGGAATGGAGATTCTACTTTGG 4506  
DB 1847 AAGAAGAAAGTACTCTTAATAATGCTCTACAAAGAAATTCAGAAATGAGTGAATTAAC 1906  
QY 4507 AGTTCCATTCCTCCTGAGTTTGGGTTGAGTCTTCGATTAATAATTCAAATAGATCA 4566  
DB 1907 AGCTCCCTGAAGGTTACGGCTGATTTTGGTCTTAAGATCCCTTAATCTGAGAGCTAA 1966  
QY 4567 AGGAAATGTTATCTCTTTGGCACAACAAATTTATAGAAATGGAATGAAAGTGCCACCTG 4626  
DB 1967 AAGAAATGCTAAACGACATTTGGAACCTGCTACCTACAAAGTGCGACTAAAGGTTCTATCCA 2026  
QY 4627 ATGAAGGATCCTCGATCCCATGCTGACCTGGAGCACCTGGCTTCTACTATCCAG 4686  
DB 2027 ATGAAGAGGATCCTCGTGTCCCAATAATGTTTGGGTTAGCTGCGGTACACCATCCAAA 2086  
QY 4687 CAATTGAAATCTATTGGGAGATGAAGAAACCTCTGTTTGGAGCATTCAAAATAGGC 4746  
DB 2087 GCATAGAAAGATTTTGAAGTATAGATAAACCATGTTTGGTCTTTTACCTTGCAGAC 2146  
QY 4747 AGCATAATGGTCTGAAGCATTAATGACGTTTGCAGTTGTCACAGAGGATTAACCTGCTTC 4806  
DB 2147 TGGATGACTGTCTTAGTCAATGAACGAGATTTGCGCAGCACACATGACAGTGGCATCAG 2206  
QY 4807 AGGTCTGTATACAGAAACATCTGGTTCGTTCTCTATCATGTTGTTCTTCTAACAATAAT 4866  
DB 2207 TTTCAGTGGTGCAAGGACATTTTGTAAACTTTTGTGCATCACTGGTGCCTTAATGACAGCC 2266  
QY 4867 CAGAAGATACACATGCTCTGCTATAGATCTGTTTCATGTTTGGTGGTGGTGTGT 4926  
DB 2267 ATGAGGAATCTTCATGATTAATGATATGACATGTTTCAATTAATGTTGGGCTGGTGC 2326  
QY 4927 TAGCATTTCCATCTTGTATTTGGATGACCTGTTGATCTGCAGCCTTCTTCAGTTAGTT 4986  
DB 2327 TTGCAATTTCTGGTGTGAGTG-----TCAGGATTTTCAGGATCAGCC 2371  
QY 4987 CTTCCTTAACCACTTTATCTCTCCATTTGATCAACCATGGGACACATGCTTTCAGATAC 5046  
DB 2372 TTGGCATTGGAGACCTTCAATTTTCCATCTGTTTACTATGTCACACATCATACAGATCT 2431  
QY 5047 TACTTACAGTACACAC---AGGCCTACCCCTGCTCAGGTTCAAGAGACAGTGAAGG 5103  
DB 2432 TACTTACCTCATGTACAGAGAGATGGCATGGATCAAGAAATCCCCCTTGTGAAGAG 2491  
QY 5104 CTCAATTCGGCATCTTCTTTCTTTGAGAAATTTCTCAATATATACAGTGGCTCCATTTGGT 5163  
DB 2492 ATACAGAGTCTTCTGCTTTGTATATAACACATTCACCATGATATACGGGAGTGCCTTGAAG 2551  
QY 5164 GTGATATCTCGGTGGTATTTTGGGTCTCACTGAAGAAATGGCATACCCCTTATCTTC 5223

DB 2552 AAATACCATCCGCTGGCATCTGTGGAGGAGTGTACAGCTGGAATCATGCTCTTCTGA 2611  
QY 5224 GCTGTGCTGCATTTGTTTTCCTACTAATTTTGGGTAACTCCGCTGAGGAACTGCATA 5283  
DB 2612 AGTGTGCTGCTTTATTTTTCATTAATTAATGAGTCTCTTCCCAACCGCATTTCAAG 2671  
QY 5284 CCATTTCTCAGAGAGAGATACAGTGCACTCTGTAGCTATCTATCTTTTACCTACAAAT 5343  
DB 2672 T---TCCTGGAACAAGCCATTTTGAACATTTATGTAGCTATCTTCCCTACCAACAACC 2728  
QY 5344 TGTTCCTGCTCTTCCAGGAATATTTGGATACCTGTAAGCCCTTGTCTCAGAGGTGGTG 5403  
DB 2729 TCATTTGCTCTTTTCAAGAAATATGAGTAATGAAATTCATGATTAAGTGTGGTGC 2788  
QY 5404 CAGATCTCTGCTTACTTAAACTCTTTGAAGCAAAAACACCGTGGTCACTACCTTAGAA 5463  
DB 2789 GTAAACAGTGAAGTAAAGATATCTAGAAAGGTGAAGAGATGCTATAAGATATCCAAGAG 2848  
QY 5464 AAGAAATATGTTGATAGAGCTTCTGTAGTACTATAGCTGCTCTCTGAAATCAAGCTTCTC 5523  
DB 2849 AATCTAAACAAATTAATAAACCTTCCAGAGGATTTACAGCAGCTCATTAATCAAGCATCCA 2908  
QY 5524 ATTTCCAGGTGCCCAACGCTCTGAGATGATGAGGAAAGCATCTGTCTCTCTGCTTTCT 5583  
DB 2909 ATTTCTGTGCTCCGAATCAGGTGTGATAGAGCAGAGCCCAACTCTGTGCTTGTGT 2968  
QY 5584 GTGGGCTTATATCTATGTTCTCAGAAACATTTGTGCTCAGGAAATTTGAAACGGGAAGAG 5643  
DB 2969 GCGGATCTCTGCTGCTCCAGAGTTACTGTGCTCAGACTGAACTGGNAGGGAGGATG 3028  
QY 5644 TTGGAGCTTGCAATTTTTCAGCACTTCACTGTGGAGCCGAGTCTGCAATTTTCTTAAAAA 5703  
DB 3029 TAGGAGCTTGCAACAGCTCACACTACTCTGTGGCTCTGGAGTGGGCACTTCTCTGAGAG 3088  
QY 5704 TCAGAGAAATGCCAGTGTCTGTTGAGGTTAAAGCCAGAGGCTGTGCTATCCAGCTC 5763  
DB 3089 TACGGAATGTCAAGTGTCTATTTTAGCTGGCAAAACCAAGGCTGTTTTTATCTCTCCTC 3148  
QY 5764 CTTTCTTGATGAATATGGAGAAACAGACCTCTGGCTTGAAGGGGCAACCCCTTCAAT 5823  
DB 3149 CTTACTTTGATGACTATGGGAGACCGACAGGACTCAGACGGGAAATCTTTTACAT 3208  
QY 5824 TATCTGTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACAACATGCAATATAGAAG 5883  
DB 3209 TATGCAAGAGCGGATTTCAAGAGATTTCAAGAGCTCTGGCACCAACACAGTGTACAGAGG 3268  
QY 5884 AGATTGCTAGGAGCCAGAGACTAATCAGATGTTTATTTGGATTCACTGGCA 5935  
DB 3269 AAATTTGGAATGTCACAGGAAGCCAAATCAGACACTGTTTGGCATTTGACTGGCA 3320

Search completed: July 9, 2005, 15:26:52  
Job time : 3366 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 18:43:15 ; Search time 17538 Seconds  
(without alignments)  
13690.809 Million cell updates/sec

Title: US-10-758-672A-1  
Perfect score: 6308  
Sequence: 1 gccagaattcgcagcagg.....aatttgatttggtgtttt 6308

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gse1: \*  
9: gb\_gse2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2441.6	38.7	4386	9 AY405198	AY405198 Homo sapi
2	2334.2	37.0	4386	9 AY405199	AY405199 Pan trogl
3	2096	33.2	2561	3 AK089616	AK089616 Mus muscu
4	1671.4	26.5	4247	9 AY405200	AY405200 Mus muscu
5	826.6	13.1	360	4 BM472160	BM472160 AGENCOURT
6	806.4	12.8	3641	3 BC044903	BC044903 Mus muscu
7	771.4	12.2	865	5 BQ233617	BQ233617 AGENCOURT
8	740.2	11.7	3058	3 AK078173	AK078173 Mus muscu
9	723	11.5	838	1 AJ820333	AJ820333 Mus muscu
10	710.4	11.3	712	5 BX104087	BX104087 BX104087
11	687.8	10.9	3627	3 AK083320	AK083320 Mus muscu
12	674.6	10.7	682	1 AI929033	AI929033 au64c10.y
13	668.8	10.6	898	5 BU155953	BU155953 AGENCOURT
14	664.8	10.5	668	5 BU618516	BU618516 UI-H-FH1-
15	656.6	10.4	756	4 BG534574	BG534574 602553425
16	637.4	10.1	641	1 AI361043	AI361043 QY03511.x
17	635.2	10.1	774	5 BU955331	BU955331 AGENCOURT
18	627.2	9.9	797	4 BG862813	BG862813 602799074
19	600	9.5	644	7 CN276943	CN276943 170005321
20	594.8	9.4	640	6 CD368666	CD368666 UI-H-FH1-
21	585.2	9.3	695	6 BY756516	BY756516 BY756516
22	583.6	9.3	703	4 BM114217	BM114217 L0801D04-
23	576	9.1	729	5 BU703363	BU703363 UI-M-PD0-
24	573.2	9.1	782	4 BI086469	BI086469 602849734

25	567.4	9.0	689	7 CN528381	CN528381 UI-M-HQ0-
26	560	8.9	560	6 CB159971	CB159971 K-EST0219
27	557.4	8.8	681	7 CK771118	CK771118 959357 WA
28	557	8.8	565	2 BF063405	BF063405 Th89d06.x
29	553.4	8.8	567	7 CN276944	CN276944 170004241
30	553.2	8.8	730	7 CV116540	CV116540 AGENCOURT
31	551.4	8.7	778	7 CF741991	CF741991 UI-M-HB0-
32	545.4	8.6	684	6 BY764809	BY764809 BY764809
33	539	8.5	751	7 CF726597	CF726597 UI-M-HB0-
34	536	8.5	774	7 CK635421	CK635421 UI-M-HN0-
35	533.4	8.5	744	7 CF539210	CF539210 UI-M-GH0-
36	529.2	8.4	691	6 BY736138	BY736138 BY736138
37	521.2	8.3	528	5 BX119443	BX119443 BX119443
38	517.6	8.2	659	6 BY750548	BY750548 BY750548
39	512.2	8.1	606	2 BE589438	BE589438 195602 BA
40	511.4	8.1	752	6 CA313244	CA313244 UI-CF-FN0
41	509.4	8.1	633	2 BB622408	BB622408 BB622408
42	501.4	7.9	506	5 BX486919	BX486919 DKFZ6686B
43	499.6	7.9	777	6 CB232501	CB232501 AGENCOURT
44	497.4	7.9	719	5 BU623096	BU623096 UI-H-FL1-
45	496	7.9	496	5 BX281065	BX281065 BX281065

## ALIGNMENTS

RESULT 1  
AY405198  
LOCUS AY405198 4386 bp DNA linear GSS 13-DEC-2003  
DEFINITION Homo sapiens HCM2144 gene, VIRTUAL TRANSCRIPT, partial sequence,  
Genomic survey sequence.  
ACCESSION AY405198  
VERSION AY405198.1 GI:39761172  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4386)  
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B.,  
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 4386)  
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B.,  
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
Location/Qualifiers  
source 1..4386  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>4386  
/locus\_tag="HCM2144"

Query Match 38.7%; Score 2441.6; DB 9; Length 4386;  
Best Local Similarity 65.2%; Pred. No. 0;  
Matches 3440; Conservative 0; Mismatches 916; Indels 924; Gaps 6;  
QY 696 ATGGCGGACGAGCGCTGGAGGTACTGAGAGATGGAATCAGCGCGAGTTACCCAG 755  
|||||  
DB 1 ATGGCGGACGAGCGCTGGAGGTACTGAGAGATGGAATCAGCGCGAGTTACCCAG 60

Qy	756	ACCCCTCAGCGTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATATCTGCTTCTTG	815
Db	61	ACCCCTCAGCGTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATATCTGCTTCTTG	120
Qy	816	CATCATTTGGCACAATTTGGTGCAGAAATTTACTTGTCTGAAATGGACCCAGACTTGGAA	875
Db	121	CATCATTTGGCACAATTTGGTGCAGAAATTTACTTGTCTGAAATGGACCCAGACTTGGAA	180
Qy	876	AAGCAGGAGGAAAGTCTACAAATGTCAATATTTCACTTCCACTGGAATGTACTTATTTGGA	935
Db	181	AAGCAGGAGGAAAGTCTACAAATGTCAATATTTCACTTCCACTGGAATGTACTTATTTGGA	240
Qy	936	GAAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCAGTGGAGCAATTTCAAGTTTGTGGG	995
Db	241	GAAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCAGTGGAGCAATTTCAAGTTTGTGGG	300
Qy	996	AGGGTTTTCAAAAGTGGAGAGACAACCTATTTCTTGAGGGATTTGCAATTTGATCCACA	1055
Db	301	AGGGTTTTCAAAAGTGGAGAGACAACCTATTTCTTGAGGGATTTGCAATTTGATCCACA	360
Qy	1056	TGTGTACTCTGTATGACTCTTCCAGGACAGTGTTCATAAATTCATCGTTTACAGATG	1115
Db	361	TGTGTACTCTGTATGACTCTTCCAGGACAGTGTTCATAAATTCATCGTTTACAGATG	420
Qy	1116	CATACCTTCTACTCGAGAGGGTTCTGTACTGTGGAGACACAGAGGCTAGAAATCTGGC	1175
Db	421	CATACCTTCTACTCGAGAGGGTTCTGTACTGTGGAGACACAGAGGCTAGAAATCTGGC	480
Qy	1176	CTTTTTTGTGTAATCATGCACTGGAAGCAGGAGTACTATATAAGAGAAATTCACGCTGT	1235
Db	481	CTTTTTTGTGTAATCATGCACTGGAAGCAGGAGTACTATATAAGAGAAATTCACGCTGT	540
Qy	1236	CCGTTGAATGAAGAGTAAATTTGTCGAAGCCAGGAAATATTTCTTTCAGTGATAAATAT	1295
Db	541	CCGTTGAATGAAGAGTAAATTTGTCGAAGCCAGGAAATATTTCTTTCAGTGATAAATAT	600
Qy	1296	GTCTGAGAAATGACTATATGGGAAGAGGAAAGAACTGCCCTCTGAACTCCAGATAAGG	1355
Db	601	GTCTGAGAAATGACTATATGGGAAGAGGAAAGAACTGCCCTCTGAACTCCAGATAAGG	660
Qy	1356	GAGAAAATGAAAGATACCTATTTGTGTCTTTTCAATGATGAACACATTCATATGACAC	1415
Db	661	GAGAAAATGAAAGATACCTATTTGTGTCTTTTCAATGATGAACACATTCATATGACAC	720
Qy	1416	GTCTATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTGCATACC	1475
Db	721	GTCTATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTGCATACC	780
Qy	1476	ACTGCCATTCACAAAGAGGCTGTGGGCTGTTTAAAGCGGAGCTTATGTCTTGCCAG	1535
Db	781	ACTGCCATTCACAAAGAGGCTGTGGGCTGTTTAAAGCGGAGCTTATGTCTTGCCAG	840
Qy	1536	GAAGCAAGGAAGATATAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACTTCATGTA	1595
Db	841	GAAGCAAGGAAGATATAAGAGTCAATTCAGAAAATGTCTCTCAACATCCACTTCATGTA	900
Qy	1596	GAAGTATTAACCTCAGAGATTAATGGCTCATCAGAAAATTTGCTTGGGTCCTGG	1655
Db	901	GAAGTATTAACCTCAGAGATTAATGGCTCATCAGAAAATTTGCTTGGGTCCTGG	960
Qy	1656	ATGACAAATTTATGACTTCAAGTGAATTTAGGCAGATCTTTTGCCAGCATGCCCTT	1715
Db	961	ATGACAAATTTATGACTTCAAGTGAATTTAGGCAGATCTTTTGCCAGCATGCCCTT	1020
Qy	1716	AGAGAAGAACTGACTCGAGAAATCCCTGTCTCATTAAGCAGAGTAAATGCTTTGGGATGCA	1775
Db	1021	AGAGAAGAACTGACTCGAGAAATCCCTGTCTCATTAAGCAGAGTAAATGCTTTGGGATGCA	1080
Qy	1776	AAGCTTTTAAAGTGGCCGTAAGATCCCTTCATGAATTCATCTTCAGCAGTTTTTTTATG	1835
Db	1081	AAGCTTTTAAAGTGGCCGTAAGATCCCTTCATGAATTCATCTTCAGCAGTTTTTTTATG	1140

Db 2221 |||||GATTTGATTTAAACAAATATATATACACTAAATAGAGAAATGCTTCAGGTCTCTCATATATT 2280  
QY 2946 GTGGGTGAGCGTTATGTACCTGGAGTGGGAATGTGACCAAGAAGAGGTCACAATGAGA 3005  
Db 2281 |||  
GTG----- 2283  
QY 3006 GAAATCAITTCACITTCCTTGGCATTGAAACCATGCCACACAGTGCATTTGCCAAAATTTA 3065  
Db 2284 ----- 2283  
QY 3066 CCTGAGAAATGAAATAATGAAACTGGCTTAGAGAAATGCTAATAACAAGTGGCCACATTT 3125  
Db 2284 ----- 2283  
QY 3126 AAGAAACAGGTGTATCAGGCCATGGAGTTTATGAACATAAAGATGAATCACTGAAAGAC 3185  
Db 2284 ----- 2283  
QY 3186 TTCAATATGTACTTTTATCATTTACTCTCAAAACCCAGCATAGCAGGCTGAACATATGCAG 3245  
Db 2284 ----- 2283  
QY 3246 AAGAAAGGAGAAACCAAGAAACAAAGATGAGCAATTCGCCACACACCTCTGAA 3305  
Db 2284 -----GCATTCGCCGCCACACACCTCTGAA 2310  
QY 3306 TTCTGCCCTGCTTTTCAGCAAAGTGATTAACCTTCTCAACTGTGATATCATGTATGATTA 3365  
Db 2311 TTCTGCCCTGCTTTTCAGCAAAGTGATTAACCTTCTCAACTGTGATATCATGTATGATTA 2370  
QY 3366 CTCAGAACCGTATTTGAGCGGGCAATAGACAGATTCCTAACTGTGGAACCGAAGGGATG 3425  
Db 2371 CTCAGAACCGTATTTGAGCGGGCAATAGACAGATTCCTAACTGTGGAACCGAAGGGATG 2430  
QY 3426 CTCCAAATGGCTTTTCATATCTGGCATTTGGGTTTACTAGAGAGAGCAACAGCTTCAA 3485  
Db 2431 CTCCAAATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2490  
QY 3486 AAAGCTCTCTGAAGAAGATPAACATTTGACTTTTATCATAGGCTTCAAGATTTGGGAAGT 3545  
Db 2491 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2550  
QY 3546 TCAGGCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAATCCCCAGTTAGAAGGC 3605  
Db 2551 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2610  
QY 3606 CAGAAGGACATGATAACGTGGATACCTTCAGATGTTTGAACACAGTGAAGCGATTAGAGAA 3665  
Db 2611 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2670  
QY 3666 AAATCTTGTTTAAATTTGTAGCAACCATCAGGATCGGAATCTATTAAAGAAATGATGAGATT 3725  
Db 2671 AAATCTTGTTTAAATTTGTAGCAACCATCAGGATCGGAATCTATTAAAGAAATGATGAGATT 2730  
QY 3726 ACTCATGATAAAGAAAAAGCAGAACGAAAGAAAGCTGAAGCTGTAGGCTACATGCG 3785  
Db 2731 ACTCATGATAAAGAAAAAGCAGAACGAAAGAAAGCTGAAGCTGTAGGCTACATGCG 2790  
QY 3786 CAGAAGATCATGGCTCAGATGTCTGCTTTACAGAAAACTTCATTGAAAACTCATAAATCTC 3845  
Db 2791 CAGAAGATCATGGCTCAGATGTCTGCTTTACAGAAAACTTCATTGAAAACTCATAAATCTC 2850  
QY 3846 ATGTATGACAAATACATCAAAATGCTTGGGAAGAGATTTCCATTATGGAGGAGAGAGC 3905  
Db 2851 ATGTATGACAAATACATCAAAATGCTTGGGAAGAGATTTCCANNNNNNNNNNNNNNNN 2910  
QY 3906 ACCCAGCAGTCAGTGACTACTCTAGAAATTTGGGTCTTAAACGGGGTCCATCTGTT 3965  
Db 2911 ACCCAGCAGTCAGTGACTACTCTAGAAATTTGGGTCTTAAACGGGGTCCATCTGTT 2970  
QY 3966 ACTGAAAGGAGGTCTGACGTGCATCTTTGCCAAGAGAAACAGAGGTGAAAAATAGAA 4025  
|||||

2971 ACTGAAAGGAGGTGCTGACGTGCATCTTTGCCAAGAGAAACAGAGGTGAAAAATAGAA 3030  
QY 4026 AATAATGCCATGATTAATTCGGCTGTGTCCAGAAATCTACTGCTTTAAACCCAGCACAGG 4085  
Db 3031 AATAATGCCATGATTAATTCGGCTGTGTCCAGAAATCTACTGCTTTAAACCCAGCACAGG 3090  
QY 4086 GGAATAACCATAGAACTCTCAGGAGAGCCCTAGACCCACTTTTCATGGATCCAGACTTG 4145  
Db 3091 GNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3150  
QY 4146 GCATATGGAACCTTATACAGGAAGCTGTGTCAATGTAATGACGACAGTGTGTGCGAGAAG 4205  
Db 3151 GCATATGGAACCTTATACAGGAAGCTGTGTCAATGTAATGACGACAGTGTGTGCGAGAAG 3210  
QY 4206 TATTTTGAAGCTGTACAGCTGAGCTCTCAGACGCGCAATCATGTGACCTTTTGTGACTTG 4265  
Db 3211 TATTTTGAAGCTGTACAGCTGAGCTCTCAGACGCGCAATCATGTGACCTTTTGTGACTTG 3270  
QY 4266 GAAAGTGGAGAAATATCTTTGCCCTCTTTGCAATCTCTGTGCAATCTGTGATCCCATTT 4325  
Db 3271 GAAAGTGGAGAAATATCTTTGCCCTCTTTGCAATCTCTGTGCAATCTGTGATCCCATTT 3330  
QY 4326 ATTTCTTTTGCACAACTCAAAAGATATAACAGTGAGATGACAGTCTCTCTCAACTTTTG 4385  
Db 3331 ATTTCTTTTGCACAACTCA----- 3347  
QY 4386 ACCCTGGCAGCGTGGATACAGACTGTTCTGGCCAGAAATACAGTTTATAATATAAGACAT 4445  
Db 3348 ----- 3347  
QY 4446 GCTAAGGAGAAACCCAAATTCCTATTTCTTTAAATCAAGGAATGGGAGATTCTACTTTG 4505  
Db 3348 ----- 3347  
QY 4506 GAGTTCCATTCCATCCTGAGTTTGGCGTTGAGTCTTCGATTAAATATTCAATAGCATC 4565  
Db 3348 -----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3381  
QY 4566 AAGGAAATGGTTATTTCTCTTTGCCACAACTTTATAGAAATTTGGAATTTGAAAGTGCCACCT 4625  
Db 3382 AAGGAAATGGTTATTTCTCTTTGCCACAACTTTATAGAAATTTGGAATTTGAAAGTGCCACCT 3441  
QY 4626 GATGAAAGGATCTCTGAGTCCCAATGCTGACCTGGAGCACCTGGCTTTTCAATATCCAG 4685  
Db 3442 GATGAAAGGATCTCTGAGTCCCAATGCTGACCTGGAGCACCTGGCTTTTCAATATCCAG 3501  
QY 4686 GCAATTTGAAATCTATTTGGGAGATGAAGGAAACCTCTGTTTGGGAGCACTTCAAAATAGG 4745  
Db 3502 GCAATTT-----GGTAAAGCTCATAGAAACCTGGTTTGGAGGTT 3540  
QY 4746 CAGCATAATGGTCTGAAAGCATTAAATGCAATTTGAGTTTGCAGTTTGCACAGAGGATTACCTGTCT 4805  
Db 3541 TCACATAATGGTCTGAAAGCATTAAATGCAATTTGAGTTTGCAGTTTGCACAGAGGATTACCTGTCT 3600  
QY 4806 CAGGTCTCTGATACAGAAACATCTGTTTCTTCTATCAGTTGTTTCTTCTTAACATAAAA 4865  
Db 3601 CAGGTCTCTGATACAGAAACATCTGTTTCTTCTAT----- 3637  
QY 4866 TCAGAGATACACCACTGCTCTGTATAGATCTGTTTCTATAGATCTGTTTGTGGGTGCTGTG 4925  
Db 3638 ----- 3637  
QY 4926 TTGATCTTCCCATCTCTGTATTTGGATGACCTGTTGATCTGCAGCCTTCTTCAGTTAGT 4985  
Db 3638 ----- 3637  
QY 4986 TCTTCTATAACCACTTTATCTCTTCCATTTGATCACAATGGCACACATGCTTCAGATA 5045  
Db 3638 ----- 3637  
QY 5046 CTACTTACAGTAGACAGAGGCTACCCCTTGTCTGAGTTCAAGAGACAGATGAGGGCT 5105  
Db 3638 -----CAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3681



1356 GAGAAAAATGAAAGATACTATTGCTGCTCTTTCAATGATGAACACCACTTCATATGACCAC 1415  
1357 |||||  
661 GAGAAAAATGAAAGATACTATTGCTGCTCTTTCAATGATGAACACCACTTCATATGACCAC 720  
662 |||||  
1416 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTGCGAGAGCCCAAGTTGCATACC 1475  
1417 |||||  
721 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTGCGAGAGCCCAAGTTGCATACC 780  
722 |||||  
1476 ACTGCCATTGACAAAGAGGCTGTCGGGCTGTTAAAGCGGAGCTTATGCTGCTTCCGAC 1535  
1477 |||||  
781 ACTGCCATTGACAAAGAGGCTGTCGGGCTGTTAAAGCGGAGCTTATGCTGCTTCCGAC 840  
782 |||||  
1536 GAAGCAAAAGAGATATAAAGAGCTCAATTCAGAAAATGCTCTCAACATCCACTTCATGTA 1595  
1537 |||||  
841 GAAGCAAAAGAGATATAAAGAGCTCAATTCAGAAAATGCTCTCAACATCCACTTCATGTA 900  
842 |||||  
1596 GAAGTATTACACTCAGAGATTATGGCTCATCAGAAAATTTGCTTTGGCTCTTGCTTCCCTG 1655  
1597 |||||  
901 GAAGTATTACACTCAGAGATTATGGCTCATCAGAAAATTTGCTTTACGCTCTGGTTCTCG 960  
902 |||||  
1656 ATGAACAAAATTTATGAGCTATTCAAGTGACTTTAGGCAGATCTTTTGCAGCATGCCCTT 1715  
1657 |||||  
961 ATGAACAAAATTTATGAGCTATTCAAGTGACTTTAGGCAGATCTTTTGCAGCATGCCCTT 1020  
962 |||||  
1716 AGAGAGAACCTGACTCGAGATCCCTGCTCATTAAGCAGGTAATGCTTTGGGATGCA 1775  
1717 |||||  
1021 AGAGAGAACCTGACTCGAGATCCCTGCTCATTAAGCAGGTAATGCTTTGGGATGCA 1080  
1022 |||||  
1776 AAGCTTTATAAAGGTCGCCGTAAGATCTTCTCATGAATTCATCTTCAGCAGTTTCTTTATG 1835  
1777 |||||  
1081 AAGCTTTATAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140  
1082 |||||  
1836 GAGATGGAATAAAAAAATCTTTTGTGATGGAATTTGTGAAGTATTTATAAACAACATGCGAG 1895  
1837 |||||  
1141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1200  
1142 |||||  
1896 AAGAAATATATCAGTGATGATCATGACAGAGTATCTCTATACTGCACTTTTCAGTTGAG 1955  
1897 |||||  
1201 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1260  
1202 |||||  
1956 ATGTTTACTGTTCTACTCTGCTCGACATCTTATTGAAGAGCAGATGTTATCTCTGTC 2015  
1957 |||||  
1261 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1320  
1262 |||||  
2016 ATTACTGAAACTCTGCTAGAAGTTTACCTGAGTACTTGGACAGGAACAATAAATTCAC 2075  
2017 |||||  
1321 ATTACTGAAACTCTGCTAGAAGTTTACCTGAGTACTTGGACAGGAACAATAAATTCAC 1380  
1322 |||||  
2076 TTCCAGGGTTATAGCCAGGACAAAATTTGGGAAGAGTATATGCAAGTAAATATGTCACCTAAAG 2135  
2077 |||||  
1381 TTCCAGGGTTATAGCCAGGACAAAATTTGGGAAGAGTATATGCAAGTAAATATGTCACCTAAAN 1440  
1382 |||||  
2136 TATATCTCTGATCAGCAAAACCCACAATATGCAAGAAAGATTAAAGATGCAAGTTCTCTGAA 2195  
2137 |||||  
1441 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1500  
1442 |||||  
2196 GGTGTTTCGATCTTTTGAAGATTTCTTACCTGATGAGGGAATGGAAGAAATCCGAAGA 2255  
2197 |||||  
1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1560  
1502 |||||  
2256 CAGGTTGGGCAACACATTGAAGTGAATCTGATGAGGAGCTGCCATGCTATACAGATG 2315  
2257 |||||  
1561 CAGGTTGGGCAACACATTGAAGTGAATCTGATGAGGAGCTGCCATGCTATACAGATG 1620  
1562 |||||  
2316 CAATTTGAAGAAATTTTACTCATGTTCCAGAGTGGTGGCTGCTGATGATGAAGAACTCTTA 2375  
2317 |||||  
1621 CAATTTGAAGAAATTTTACTCATGTTCCAGAGTGGTGGCTGCTGATGATGAAGAACTCTTA 1680  
1622 |||||  
2376 CTGTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGTGAGTGACGATTTTCATATCT 2435  
2377 |||||  
1681 NNNNNNNNTTATAAAGAAATGTCACAAAGCTGTGATGAGTGAGTGACGATTTTCATATCT 1740  
1682 |||||  
2436 AGTAGCAAGACAGTAGTAGTAATCGTGTGGACATAGTTTGGAAAAAAGTCTCTACAGAGTA 2495

1741 AGTAGCAAGACAGTAGTAGTAATCGTGTGGACATAGTTTGGAAAAAAGTCTCTACAGATA 1800  
1742 |||||  
2496 TCTGAGGATCTTTGTAAGCATACATCTGCGCACTCTCTAGGACCCCTGCTGCTTCATGTA 2555  
2497 |||||  
1801 TCTGAGGATCTTTGTAAGCATACATCTGCGCACTCTCTAGGACCCCTGCTGNNNNNNNNNN 1860  
1802 |||||  
2556 CGTTTAAAGCAGGCTGGTCTGCTTTCAAGCTGCAATGAAATTTGCTGCTTTTGAAGACTTT 2615  
2557 |||||  
1861 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1920  
1862 |||||  
2616 CAAGTAGAGGTAAGTGGAAATATCTCTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2675  
2617 |||||  
1921 NNNNNNNNGTACTAGTGGAAATATCTCTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
1922 |||||  
2676 GCTGAGATGTCGCGAAGAAATGGAATGCTCTCTTATTAGCCAGGCTGTTTATTACCAAGAT 2735  
2677 |||||  
1981 GCTGAGATGTCGCGAAGAAATGGAATGCTCTCTTATTANNNGTGTGTTTATTACCAAGAT 2040  
1982 |||||  
2736 GTTAAAGTCAGAGAAAGAAATGATGATAAAGATATCATCATGCTT----- 2780  
2737 |||||  
2041 GTTAAAGTCAGAGAAAGAAATGATGATAAAGATATCATCATGCTTCAGNNNNNNNTATGAT 2100  
2042 |||||  
2781 -----CAGATTGGTGCATCTTTAAATGGATCCCAATAAGTTCTTGTACTG 2825  
2782 |||||  
2101 GCTGGAATTTCTTTCACAGATTGGTGCATCTTTNNNNNNNNNNNNNNNNNNNNNNNNNN 2160  
2102 |||||  
2826 GTACTTCAGAGGTATGAACTTGCAGGCTTTTAAACAGACCATATCTCAAAAGACCAG 2885  
2827 |||||  
2161 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2220  
2162 |||||  
2886 GATTGATTAATAATAATAATACTAATAAGAAATGCTTCAGTCTCTCATCTATATTT 2945  
2887 |||||  
2221 GATTGATTAATAATAATAATACTAATAAGAAATGCTTCAGTCTCTCATCTATATTT 2280  
2222 |||||  
2946 GTGGGTGAGCGTTATGTAAGTGGGAAATGTAAGAAAGAGAGTCAATGAGA 3005  
2947 |||||  
2281 GTG----- 2283  
3006 GAAATCAATTCACCTGCTGCTGATGAAACCAACACAGTGCCATTCGCCAAAAATTTA 3065  
3007 |||||  
2284 ----- 2283  
3066 COTGAGATGAAATAATAATACTGCTTGAAGAAATGTCTAATAAACAAGTGCCACATTT 3125  
3067 |||||  
2284 ----- 2283  
3126 AAGAAACCAAGGTGATCAGGCCATGAGTTTATGAATTAAGATGAATCACTGAAGAC 3185  
3127 |||||  
2284 ----- 2283  
3186 TTCAATATGTACTTTTATCATTTCTCAAAACCCAGCATAGCAAGGCTGAACATATGACAG 3245  
3187 |||||  
2284 ----- 2283  
3246 AAGAAAGGAGAAAAACAAGAAAAACAAGATGAAGCAITTCGCCACCAACCACTCTCTGAA 3305  
3247 |||||  
2284 -----GCATTTGCCGCCACCACTCTCTGAA 2310  
3306 TTCTGCCCTGCTTTTACGAGAAAGTGAATTAACCTTCTCAACTGTGATATCATGATGATACAT 3365  
3307 |||||  
2311 TTCTGCCCTGCTTTTACGAGAAAGTGAATTAACCTTCTCAACTGTGATATCATGATGATACAT 2370  
2312 |||||  
3366 CTGAGGACCGTATTTGAGCGGCAATACACAGATTTCTAACTTGTGACCCCAAGGGATG 3425  
3367 |||||  
2371 CTCAGGACCGTATTTGAGCGGCAATACACAGATTTCTAACTTGTGACCTGAAGGGATG 2430  
2372 |||||  
3426 CTCAAAATGGCTTTTTCATATTTCTGCAATGCGGTTTACTAGAAGAGAACACAGCTTCAA 3485  
3427 |||||  
2431 CTCAAAATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2490  
2432 |||||  
3486 AAGCTCTGGAAGAAAGTAAATTCATGCTTTTATCAATAGGCTTCAAGATTGGGAAGT 3545





QY 5706 AGAGATGCCAGTGGTCTGGTTGAAGTAAAGCCAGAGGCTGTGCTTATCCAGCTCT 5765

Db 4147 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4206

QY 5766 TACTTGGATGAATATGAGAAACAGACCTGGCTGGAAGAGGGCAACCCCTTCATT 5825

Db 4207 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4266

QY 5826 TCTCGTAGCGGTATCGGAAGCTCCATTGGTCTGGCAACAACACTGCATTATAGAAG 5885

Db 4267 TCTCGTAGCGGTATCGGAAGCTCCATTGGTCTGGCAACAACACTGCATTATAGAAG 4326

QY 5886 ATTGCTAGAGGCAAGACTATCATGATTTATTTGGATTCACTGGCAGTTACTGTGA 5945

Db 4327 ATTGCTAGAGGCAAGACTATCATGATTTATTTGGATTCACTGGCAGTTACTGTGA 4386

RESULT 3

AK089616

LOCUS

DEFINITION

AK089616 2561 bp mRNA linear HTC 03-APR-2004

Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:F830005C07 product:ubiquitin protein ligase E3 component n-recogin 1, full insert sequence.

ACCESSION

AK089616.1 GI:26354612

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20493374

PUBMED

11042159

REFERENCE

2

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE

20530913

PUBMED

11076861

REFERENCE

4

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409, 685-690 (2001)

MEDLINE

11076861

PUBMED

11076861

REFERENCE

5

AUTHORS

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

Nature 420, 563-573 (2002)

MEDLINE

12000000

PUBMED

12000000

REFERENCE

6

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashi, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

FEATURES

Location/Qualifiers

1..2561

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="NOD"

/db\_xref="PANTOM DB:F830005C07"

/db\_xref="taxon:10090"

/clone="F830005C07"

/tissue\_type="activated spleen"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

14..2560

/note="unnamed protein product; putative ubiquitin protein ligase E3 component n-recogin 1 (MGD) [GI:1277977, GB|NM\_009461, evidence: BLASTN, 100%, match=2560]"

/codon\_start=1

/protein\_id="BAC40933.1"

/db\_xref="GI:26354613"

/translation="MADREMDCARMDVSPPEPLAPORPASWDDQVDFVTAFLHLA QLVPIYAEADMPDLEKSESVQMSILIPLEWYLFGEEDPDICLEKLKHSQACQLCGKV PFCVDHBFGRAGTTKESLHCPNEEVIAQARRIPFVVIYVEMTWESEKELPPELQ IREKNRYCVLFNDEHSHYDHYVLSQALDCLAEALQHTTATIDKEGRRAVKAQVY ATCQAKEDIKSHENSVQHPHVLVSVMAHOKFALRLGSMWNKIMSYSSDDPQI PCQACLVEPGESENCLISRLMLWDKLYKGAHTLHLLIFSPFMWEMKCLPANEFP VARYLQKQYISDDHRSISITALSQVMTVPTLRLHLIEBQNSIVISITELVLP EYLDNNKNFQGYSDQKLGVRVAVICDLKYLISKPVITELRKAQFEGFRSLKYL LTCQMEBIRRVQGHIEVDPDWEAAITAIQMLKNILMFQEWCAQDGLLVAIKE CHKAVMRCSTNFMFSKTVOLGHSLSKYSVSEDLVSIHLPLSRLTLAGHLVLSR LGAI SRLHFEPFDSFOVEVLVYPLRLCLVAQVVAEMRRNGLSLSQVFFYQDVK CREEMVDLIMLQIGASIMDPNKFLVLRQYELTDAFNKTSIKDQDLIKQNTLLI EMLVLIYIVGRTVPGVNVIRREVMREITHLCIEPMHSAIARMLPENNET GLENVINKVATPKPKFVSGHGYELKDSLSLDFNMFYHYSKTQHSKAEHM"

CDS

Query Match 33.2%; Score 2096; DB 3; Length 2561;

Best Local Similarity 88.7%; Pred. No. 0;

Matches 2270; Conservative 0; Mismatches 290; Indels 0; Gaps 0;

QY 684 GTTTCCTCAAGATCGCGACGAGGAGGCTGAGGAGTCTGAGAGATGGAATCAGCGG 743

Db 2 GTTTCCTTAAGATCGCGACGAGGATGAGGAGGCGCGCGAGGATGACGTCAGCCG 61

QY 744 GAGTTACCCAGACCCCTCAGCGTCTGCGATCTTGGTGGGATCAGCAAGTTGATTTAT 803

Db 62 GAGCTCCCTCGCCCGCAGCGCGCATCGTGTGGGATCAGCAAGTTGATTTCTAT 121  
Qy 804 ACTGCTTTCTTGCAATCAATTTGGCAACAATTTGGTGGCCAGAAAATTTACTTTGCTGCAATGGAC 863  
Db 122 ACTGCTTTCTTACATCAATTTGGCAACAATTTAGTGGCCAGAAAATTTATTTTGTGCTGAGATGGAC 181  
Qy 864 CCAGACTTTGGAAAACGACGAGGAAAGTGTACAAATGTCAATATTCACATCCCACTGCAATGG 923  
Db 182 CAGAAATTTGGAAAAGCAAGAGAGTGTACAGATGTCAATACTCACTCTTTTGGAGTGG 241  
Qy 924 TACTTATTTGGAGAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCAACAGTGGAGCATTT 983  
Db 242 TACTTATTTGGAGAGATCCGATATTTGCTTTAGAGAAATTTAAACACAGTGGAGCGTTTC 301  
Qy 984 CAGCTTTTGGAGGGTTTTCAAAAGTGGAGAGACAACCTATTTCTTTGCGAGGATTTGTGCA 1043  
Db 302 CAGTTGTGTGGGAAGTTTTCAAAAGTGGAGAAACAACATATTCCTGTAGGGATTTGTGCA 361  
Qy 1044 ATTGATCCAAACATGTGTACTCTGTATGGAATGCTTCCAGGACAGTGTTCATAAATATCAT 1103  
Db 362 ATTGATCCAAACGTGTGTCTGTATGGAATGCTTCCAAAGTAGTGTTCATAAATATCAT 421  
Qy 1104 CGTTACAAGATGCATACCTTACTCGAGGAGGGTTCTGTGACTGTGGAGACACAGAGGCA 1163  
Db 422 CGTTACAAGATGCATACCTTACTCGAGGAGGGTTCTGTGACTGTGGAGACACAGAGCG 481  
Qy 1164 TGGAAAACCTGGCCCTTTTGTGTAAATCATGAACCTGGAAAGCAGGAGTACTATTAAGAG 1223  
Db 482 TGGAAAACCTGGCCCTTTTGTGTGAATCACAGAGCTGGAAAGCAGGAGTACTACAAAAGAG 541  
Qy 1224 AATTCACGCTGTCGTTGAATCAAGAGGTAATTTGTCAGGACAGAAATATTTCCCTCA 1283  
Db 542 AGCTTACATTTGCCATTTGAATGAAGAGTGTATGCTCAAGCCAGAGAAATATTTCCCTTCG 601  
Qy 1284 GTGATAAATATGTCTAGAAATGACTATATGGAGAGGAAAAGAACTGCTCTCTGAA 1343  
Db 602 GTGATAAATATCATTTGTAGAAATGACTATATGGAAAGAGAAAGGAATTTGCTCTCTGAA 661  
Qy 1344 CTCCAGATAAGGAGAGAAAATGAAGATATATTTGTCTCTTTCAATGATGAACACCAT 1403  
Db 662 CTCCAGATAAGGAGAGAAAATGAACGATATATTTGTCTCTTTCAACGATGAGCACCAT 721  
Qy 1404 TCATATGACCAAGTATATACAGCCATCAAGAGCTCTTGACTGTGAGCTCCAGAGGCC 1463  
Db 722 TCGTATGATCATTTGATCTACAGTCTGCAGAGAGCTCTAGATTTGCGAGCTTCGAGAGGCA 781  
Qy 1464 CAGTTGCTATACCACTGCCATTCGACAAAGAGGGTCTGCGGCTGTTAAAGCGGAGCTTAT 1523  
Db 782 CAGCTGCACACGACTGCCATCGACAAAGAGGGTTCGCGGCTGTCAAGCAGGTGTGTAT 841  
Qy 1524 GCTGCTTCAGGAAGCAAGAGAGATATAAGAGTCAATTCAGAAATGTCTCTCAACAT 1583  
Db 842 GCCACTTGCAGGAAGCAAGAGAGATAAAGAGTCACTCAGAGAAACGCTCTCAGCAC 901  
Qy 1584 CCACCTCATGTAGAGTATTAACCTCAGAGATTTATGGCTCATCAGAAATTTGCTTTGCGT 1643  
Db 902 CCCCTCATTTGGAGTGTGCACTCCGTTGTTATGGCTCACAGAAAATTCGCTCTCGCC 961  
Qy 1644 CTTGGTCTCTGGATGAACAAAATTTATGACTATTCAGGTGACTTTTAGGAGATCTTTTTCG 1703  
Db 962 CTTGGCTCTCTGGATGAACAAAATTTATGACTATTCAGGTGACTTTTAGACAGATATTTTTCG 1021  
Qy 1704 CAGCATGTCTTAGAGAGAACTGTACTCGAGAGATCCCTGTCTCATTAAGCAGGTTAATG 1763  
Db 1022 CAGGCCCTGCTCGTAGAGAACTGGCTCTGAAAATCCCTGTCTTATTAAGCAGACTAATG 1081  
Qy 1764 CTTTGGGATGCAAGCTTTATAAGGTGCCGTGAAGTCTTCACTGAAATTTGATCTTCAGC 1823  
Db 1082 CTTTGGGATGCAAACTTTATAAGGTGCCGTGAAGTCTTCACTGAAATTTGATCTTTAGT 1141  
Qy 1824 AGTTTTTTTATGGAGATGGAATACAAAACCTCTTTGCTATGGAATTTTGTGAGATTTAT 1883  
Db 1142 AGTTTTTTTATGGAGATGGAATACAAAACCTCTTTGCTATGGAATTTTGTGAGATTTAT 1201

Qy 1884 AAAACAACCTGCAGAAAGAAATATATCAGTGTATCATATGACAGAGTATCTCTATTAACATGCA 1943  
Db 1202 AAACAACCTGCAGAAAGAAAGTATCATCAGCGACGACACAGAGAGAGCAATCTTCATAACCGCC 1261  
Qy 1944 CTTTCAGTTTCAGATGTTTCTTACTCTCTGCTCGACATCTTTATTTGAAGAGCAGAAAT 2003  
Db 1262 CTGTCGCTGCAGATGCTCACCGTCCGACCTTGGCCCGGCATCTTTATTTGAAGAGCAGAAAT 1321  
Qy 2004 GTTATCTCTGTCAATTTACTGTAACCTCTGTAGAAATTTTACCTGTAGTACTTTGGACAGGAAAC 2063  
Db 1322 GTTATTTCTGTCAATTTACTGTAACCGTCTGTAGAAATTTTACCTGTATATCTTTGGACAGGAAAC 1381  
Qy 2064 AATAAATTTCAACTTCCAGGGTTATAGCCAGACAAAATTTGSGAAGAGTATATGACGTAAATA 2123  
Db 1382 AATAAATTTCAACTTCCAGGGTTATAGCCAGGACAACTGTGGAAAGAGTCTATCGCAGTTATA 1441  
Qy 2124 TGTGACCTTAAAGTATATCTCTGATCAGCAAAACCCCAATATATGACAGAAAGATTAAGAATG 2183  
Db 1442 TGTGACCTTAAAGTATATCTCTGATTAGCAAGCTGTCTATATGACAGAAACGATTTAAGAGCG 1501  
Qy 2184 CAGTTCTCTGAAAGTTTTCGATCTTTTGTGAAGATTTCTTACCTGTATGACGAGGAAATGAA 2243  
Db 1502 CAGTTCTCTGAAAGTTTTCGCTCTTTTCTGTGAAGATTTCTTACCTGTATGACGAGGAAATGAA 1561  
Qy 2244 GAAATCCCAAGACAGGTTTGGGCAACACATTTGAAGTGGATCTTGAATTTGGAGGCTGCCAT 2303  
Db 1562 GAAATCCCAAGACAGGTTTGGACAAACATTTGAAGTGGACCTTGAATTTGGAGGCTGCCATC 1621  
Qy 2304 GCTATACAGATGCAATTTGAAGAAATATTTTACTCATTTTCCAAGAGTGGTGTCTTTGTGAT 2363  
Db 1622 GCTATACAGATGCAATTTGAAGAAATATTTTGTCTCATTTTCCAAGAGTGGTGTCTTTGTGAT 1681  
Qy 2364 GAAGAACTTCTTACTTGTGGCTTTTAAAGAAATGTCAAAAAGCTGTAAATGAGGTCGATGCA 1741  
Db 1682 GAAGAACTTCTTACTTGTGGCTTTTAAAGAAATGTCAAAAAGCTGTAAATGAGGTCGATGCA 1741  
Qy 2424 AGTTTTCATATCTGTAGACAGACAGTACTCAATCTGTGTCGACATAGTTTGGAAACAAAG 2483  
Db 1742 AATTTTCATGTCGACAGTACTCAAGACAGTGTGCAATTTGTGGGTTCATAGTCTGGAAACCAA 1801  
Qy 2484 TCCTACAGATGATCTGAGGATCTTGTAAAGCATACATCTGTCCTCTCTAGGACCTTGTCT 2543  
Db 1802 TCCTACAAAGTCTCTGAGGACCTTGTAAAGCATACACCTGTCCTCTCTAGAACACTTGTCT 1861  
Qy 2544 GGTCTTCATGTACGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGCAATTTGTGTCT 2603  
Db 1862 GGTCTTCATGTACGTTTAAAGCAGACTAGGTGCTATTTCAAGACTGCAATTTGTGTCT 1921  
Qy 2604 TTTTGAAGGACTTTCAAGTAGAGTACTAGTGGAAATATCTTTTACGTTTGTCTGTGTGTGT 2663  
Db 1922 TTTTGAAGGACTTTCAAGTAGAGTACTAGTGGAGTACCCGCTGCTGCTGCTGCTGCTG 1981  
Qy 2664 GCCCAGGTTTGTGCTGAGATGTGGCGAAGAAAATGGAATGCTGCTCTTTATTTAGCCAGGTTT 2723  
Db 1982 GCTCAGGTTTGTGCTGAGATGTGGCGAAGAAACGGGCTCTCACTCATCAGCCAGGTTTTC 2041  
Qy 2724 TATTAACCAAGATGTTTAAAGTGCAGAGAGAAATGTATGATAAAGATATCATCTGCTTCA 2783  
Db 2042 TATTAACCAAGATGTTTAAATGCAAGGAGGAAATGTACGATAAAGATATCATCTGCTTCA 2101  
Qy 2784 ATTTGGTGCATCTTTTAAATGCAATCCCAATTAAGTTCTTTGTTACTGCTACTCTCAGAGGTTGAA 2843  
Db 2102 ATTTGGAGCATCTTATTAAGTATCCCAAGTCTTTGTTACTGTTACTGTTACTCTCAGAGATATGAA 2161  
Qy 2844 CTTGCGCGAGGCTTTTAAACAAGACCATATCTACAAAAGACCCAGGATTTTGAATTAACAATAT 2903  
Db 2162 CTTACTGTGCTTTTAAACAAGACCATATCCACAAAAGACCCAGGATTTTGAATTAACAATAT 2221  
Qy 2904 AATAACATTAAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGTGGAGGCTTATGTA 2963  
Db 2222 AATAACATTAAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGAGAACGTTATGTA 2281

```
QY 2964 CCTGGAGTGGAAATGTGACCAAGAGAGGTGACATGATGAGAGAAATCATTCACCTGCTT 3023
Db 2282 CCTGGAGTGGAAATGTGACCAAGAGAGGTGATTAATGAGAGAGATTAATCTCACTTACTT 2341
QY 3024 TGCATTGAACCCCATGCCACACAGTGCATTTGCCAAAAATTTTACCTGAGAAATGAATAAT 3083
Db 2342 TGCATTGAGCCCATGCCACACAGTGCATTTGCCAAAAATTTTACCTGAGAAATGAATAAT 2401
QY 3084 GAAACTGGCTTAGAGAAATGTCATAAACAAGTGGCCACATTTAAGAAACAGGTGTATCA 3143
Db 2402 GAAACTGGCTTAGAGAAATGTCATAAACAAGTGGCCACATTTAAGAAACAGGTGTATCA 2461
QY 3144 GSCCATGGCTTATGAACTAAAGATGAATCACTGAAAGACTTCAATGTACTTTTAT 3203
Db 2462 GSCCATGGCTTATGAACTAAAGATGAATCACTGAAAGACTTCAATGTACTTTTAT 2521
QY 3204 CATTTACTCCAAACCCAGCATAGCAAGCTGAAACATATGC 3243
Db 2522 CATTTACTCCAAACCCAGCATAGCAAGCTGAAACATATGC 2561

RESULT 4
AY405200
LOCUS 4247 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM2144 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY405200
VERSION AY405200.1 GI:39761174
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 4247)
AUTHORS Clark,A.G., Tanenbaum,D.M., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
2 (bases 1 to 4247)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..4247
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>4247
/locus_tag="HCM2144"

ORIGIN
Query Match 26.5%; Score 1671.4; DB 9; Length 4247;
Best Local Similarity 58.2%; Pred. No. 0;
Matches 3055; Conservative 0; Mismatches 1192; Indels 1003; Gaps 11;

QY 696 ATGGCGGACGAGAGAGCTGAGTACTGAGAGATGGAATCAGCGCGGAGTTACCCGAG 755
Db 1 ATGGCGGACGAGAGATGAGACGCGCGGAGAGATGAGCGTCAAGCGCGGAGCTCCCTCG 60
QY 756 ACCCTCTCAGCGCTGCGCATCTGCTGGATCAGCAAGTTGATTTTATCTGCTTCTTG 815
Db 61 GCCCGGACGCGCGCGCATCGTGGTGGATCAGCAAGTTGATTTTATCTGCTTCTTCTTA 120
```

```
QY 816 CATCATTTGGCCACAATTGGTGCCAGAAATTTTACTTGTCTGAAATGACCCAGACTTGGAA 875
Db 121 CATCATTTGGCCACAATTAGTGCCAGAAATTTTATTTTGTCTGAGATGGACCCAGATTTGGA 180
QY 876 AACGAGGAGAAAGTGTACAAATGTCAATATTCACCTCCACTGGAATGGTACTTATTTTGGA 935
Db 181 AACGAGGAGAGAGTGTACGATGTCAATACTCACTCTTTGGAGTGGTACTTATTTTGGA 240
QY 936 GAAGATCCAGATAATTTGCTTTAGAGAAATTTGAAGCAGAGTGGAGCATTTTACGCTTTTGGG 995
Db 241 GAGGATCCGATATTTGCTTTAGAGAAATTTAAACACAGTCGAGCGTTCCAGTTGTGTGGG 300
QY 996 AGGGTTTCAAAGTGGAGAGACAACTATTCTTGAGGGATTTGCGAATTTGATCAACA 1055
Db 301 AAGGTTTCAAAGTGGAGAGAACACATATTTCTGTAGGGATTTGCGAATTTGATCAACG 360
QY 1056 TGTGTAATCTGTATGACTGCTTCCAGGACAGTGTTCATAAAATCATCGTTACAAGATG 1115
Db 361 TGTGTAATCTGTATGACTGCTTCCAAAGTAGTGTTCATAAAACCATCGTTACAAGATG 420
QY 1116 CATACTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGGCATGGAACCTGGC 1175
Db 421 CATACTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGGCATGGAACCTGGC 480
QY 1176 CTTTTTGTGTAAATCATGAACTGGAAGACAGGTACTATATAAGAGAAATTCACGCTGT 1235
Db 481 CTTTTTGTGTGATCACGAGCTCGAAGACAGGTACTATAAAAGAGAGCTTACATTTG 540
QY 1236 CGGTGAATGAGAGGTAAATGTCTCAAGCCAGGAAATATTTCTTCAGTGATAAATAT 1295
Db 541 CCATTGAATGAGAGGTAAATGTCTCAAGCCAGGAAATATTTCTTCAGTGATAAATAT 600
QY 1296 GTCTGAGAATGACTATATGGAAGAGGAAAGAACTGCTCTCTGAACTCCAGATAAGG 1355
Db 601 ATTGTAGAAATGACTATATGGAAGAGGAAAGAAATGCTCTCTGAACTCCAGATAAGG 660
QY 1356 GAGAAATGAAAGATACATTTGTGTCTTTCAATGATGAACACCATTCATATGACCAC 1415
Db 661 GAGAAATGAAAGATACATTTGTGTCTTTCAAGATGAGACCATTCATGATGATCAT 720
QY 1416 GTCATATACAGCTTACAAAGAGCTCTTGTGACTGTGAGCTGCGAGAGCCAGTTGCATACC 1475
Db 721 GTGATCTACAGCTTGCAGAGAGCTCTAGATTTGAGCTTGCAGAGGACAGCTGCAACG 780
QY 1476 ACTGCCATTGACAAAGAGGCTGCTCGGGCTGTATAAGCGGAGCTTATGCTGCTGCCAG 1535
Db 781 ACTGCCATCGACAAAGAGGCTGCGGGCTGTCAAGCAGGTGTGTATGCCACTTGCAG 840
QY 1536 GAAGCAAGAGAGATATAAAGAGTCAATCAGAAATGTCTCTCAACATCCACTTCATGTA 1595
Db 841 GAAGCAAGAGAGATATAAAGAGTCACTCAGAGAACGCTCTCTCAGCACCCCTCCATGTG 900
QY 1596 GAAGTATTACACTCAGAGATTTATGCTCATCAGAAATTTGCTTTGCGCTTGTGTTCTTG 1655
Db 901 GAAGTGTCTCACTCCGTTGTTATGCTTCCAGAAATTTGCTTGTGCGCTTGTGCTCTG 960
QY 1656 ATGAACAAAAATTTATGAGCTATTCAAGTGACTTTTAGGCAGATCTTTTGCAGAGCATGCCCTT 1715
Db 961 ATGAACAAAAATTTATGAGCTATTCAAGTGACTTTTAGACAGATATTTTGCAGGCTGCCCTC 1020
QY 1716 AGAGAAGAACTGACTCGGAGAAATCCCTGTCTCATAGAGAGTTAATGCTTTGGGATGCA 1775
Db 1021 GTAGAAGAACTGGCTCTGAAATCCCTGTCTTATAAGCAGAGTAAATGCTTTGGGATGCA 1080
QY 1776 AAGCTTTATAAGGTCGCCGTAAGATCCTTCAATGAATTCATTCAGCAGTGTGTTTTATG 1835
Db 1081 AAACCTTTATAAAGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140
QY 1836 GAGATGGAATACAAAAAACTCTTTTGTCTATGGAATTTGTGAAGTATTTATAAACAACATG 1895
Db 1141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1200
QY 1896 AAAGAATATATCAGTGATGATCATGACAGAGTATCTCTATAACTGCACTTTTCAGTTTCAG 1955
```









Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 40 Row: b Column: 10  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22122576  
 This clone has the following problem: frame shifted.

## FEATURES

source

1. .3641  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="Czech II"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3493115"  
 /tissue\_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."  
 /clone\_lib="NCI CGAP\_lu29"  
 /lab\_host="DH10B"  
 /notes="Vector: pCMV-SPORT6"

## ORIGIN

Query Match 12.8%; Score 806.4; DB 3; Length 3641;  
 Best Local Similarity 54.7%; Pred. No. 2.2e-202;  
 Matches 1934; Conservative 0; Mismatches 1516; Indels 88; Gaps 13;

QY 2460 TGTGGACATAGTTTGGAAACAAAGTCTACAGAGTATCTGAGGATCTTGAGGATCTTGTAGCATAT 2519  
 DB 22 TGTGGACATCTGGTGGAAACCAATCAGATATCTGTGTTTCCCAAGAAAGTTAGCATTCAC 81

QY 2520 CTGCCACTCTAGGACCTTGTGCTTCTCATGTAGTCTTAAAGCAGGCTGGGTGCTGT 2579  
 DB 82 CTCCCAATTTCTGCTTGTGCTTGTGAGGTTTGCATGTATTGTTAAGCAAAAGTGAAGTGGCA 141

QY 2580 TCAAGCTGATGAATTTGTGCTTTTGGAGGACTTTCAAGTAGAGGTA CTAGTGGAAAT 2639  
 DB 142 TATAAATTTCCAGAGCTCTTACCTTAAGTGAACCTGAGCCACCCATGTTGTAGAGCAT 201

QY 2640 CCTTTAGCTGTCTGTGTTGTTGCCAGGTTGTTGCTGAGATGTGGCGAAGATGGA 2699  
 DB 202 CCTCTTAGATGCTTCTCTTATGTGCTCAAGTGCATGCTGGGATGTGGAGAAAGATGCG 261

QY 2700 CTGCTCTTATAGCAGGTGTTTATTAACCAAGATGTTAAGTGCAGAGAAATGAT 2759  
 DB 262 TTCTCTTAGTAATCAGATCTATTACTACCAATATGTGAAATGCGGCGAGAGATGTTTC 321

QY 2760 GATAAAGATATCATCATGCTTCAAGTTGGTGCACTTTTAATGGATCCCAATAGTTCTTG 2819  
 DB 322 GACAGGACATAGTGATGCTTCAAGCAGGTGCTCCATGATGATCCAAACCACTTTCCTG 381

QY 2820 TTACTGGTACTCAGAGGTATGAACTTGGGAGGCTTTTAAACAGACCATATCTACAA- 2878  
 DB 382 ATGATCATGCTCAGCGCTTTGAACTCTATCAGCTCTTCAAGCAGGCTGACTATGGGAAG 441

QY 2879 -----AGACCAGGATTTGATTAAACAATATAACACTAATAGAA 2918  
 DB 442 AGATTTCAGTTCTGAGGTTTACCCATAGGACGTCGTTACAGAGAAACAACTCTGATCGAA 501

QY 2919 GAAATGCTTCAGGCTCCTCATATATTGTTGGGTGAGCGTTATGTACTTGGAGTGGGAAT 2978  
 DB 502 GAGATGCTCTCCTCATCATGCTTGTGGGAGAAAGATTCAACCTCTGGGTGGACAG 561

QY 2979 GTGACCAAGAGAGGTGACATGAGAGAAATCATTTCACTTGTGTTGGATTGAAACCATG 3038  
 DB 562 GTGGCTGCCACAGATGAATCAAGAGGGAGATTATCCATCAGCTGAGCATCAAACTATG 621

QY 3039 CCACACAGTGCCATTGCCAAAATTTTACCTGGAATGAAATTAATGAAATCGCTAGAG 3098  
 DB 622 GCTCAGCAGCAGCTGTGAAAGTCTCTGCTCTGAAGATGAGAACAGAGACCGGCGAGAG 681.

QY 3099 AATGTCATAAACAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGAGATTTAT 3158  
 DB 682 AGCGTTCATCGAGTCCGTAGCACATTTCAAGAAACCTGGCCTCAAGAGCAGGACATGTAT 741

QY 3159 GAACTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTAATCTCCAAAC 3218  
 DB 742 GAGCTGAGCCAGAGTGTGCCAAGAGTTCACTGTATTTTATCATTTCTTCAGGGCA 801

QY 3219 CAGCATAGCAGGCTGAACATATGCAAGAAAGAGAGAAACAGAGAAACAAAGATGAA 3278  
 DB 802 GAGCAGTCCAAAGCAGGAGGAGCTCAGCGGAAATTCGAAAGAGAGAAATTAAGAGATACA 861

QY 3279 GCATTCGCCGCCACCAACCACTCTCTGAATTTCTGCCCTGCTTTCAGCAAGGTATTAACCTT 3338  
 DB 862 GCATCTCCCTCTCCGGCTTTGCCACCGTTCTGCCCTTTGTTCCGAGTCTGGTTAAATC 921

QY 3339 CTCACACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGGCAATFAGACACA 3398  
 DB 922 TTGCAGTGTGACGTCTATGCTGTACATCATGAGGACGATCTTCGAGTGGGCTGTAGAGCAT 981

QY 3399 GATTTCACTTGTGGACCGAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTGGGT 3458  
 DB 982 CACGGGTCTGCTGGTTCAGAGTCCATGCTACAGAGGGTGTGTCATTTGATCGGATGGCT 1041

QY 3459 TTACTAGAGAGAGAGCAACAGCTTCAAAAAGCTCTCTGAAGAGAGT-----AACATTTGAC 3515  
 DB 1042 CTCACGAGAGAGAGCAACCACTTGGAGATGCGTGGAGGGCATGTGCAGACCTTTCAAC 1101

QY 3516 TTTTATCATAGGCTTCAAGATTGGGAGTTTCCAG-----CCATGATATACAAATG 3566  
 DB 1102 TTCAACAGAGATTTCAGAGCTGTGTATGACCACTAATACTCCCGAGCATCTTAGCT 1161

QY 3567 CTTTGTGAAAAAATCTCAAAGAAATTTCCAGTTTAGAAGGCCAGAGGACATGATAAGCTGG 3626  
 DB 1162 ATGCTGGAGACCTTGCAGAACCGCCCTCTCCGGAAGCCCAAGAGACATGATCAGGTGG 1221

QY 3627 ATACTTCAGATGTTTGACACAGTGAAGCGATTTAAGAGAAAAATCTTTGTTTAAATGTAGCA 3686  
 DB 1222 TTGCTAAAGATGTTTAAATGCAATTAAGAGATAAGAGA-----GTGTTTCATCCAGCAGC 1275

QY 3687 ACCACATCAGGATCGGATCTTATTAAGATGATGAGTTACTCATGATTAAGAGAAAGCA 3746  
 DB 1276 CCGTGGCGGAGCGGAGGGAACCATATATGAGGAGAGAGCTCAAGAGACAGGACAAAGCA 1335

QY 3747 GAAACGAAAAAGAAAGCTGAAGCTGTAGGCTTACATCGCCAGAGATCATGGCTCAGATG 3806  
 DB 1336 GAGAGAAAAAGAAAGCGGAGATCGCAGACTGCGCGGAGAGAGATCATGSCCAGATG 1395

QY 3807 TCTGCTTTACAGAAAACTTCAATGAACTCAATAAATCATGTATGACAAATACATCAGAA 3866  
 DB 1396 TCTGAGATGCGAGCGGACCTTCAATGACGAAAAACAAAGAGCTCTTCCAGCAGACCTTAGAG 1455

QY 3867 ATGCTCGGGAAGAGATTTCCATTTAGGAGAGAGGAGGAGCCAGCAGTCTAGTACTAC 3926  
 DB 1456 CTGGAACCTCTGCTCTGCCACTCTT-----GACAGCAGCCTCTCCGCTTTTCAGACGCA 1509

QY 3927 TCTAGAAATGCTTTTGGGCTCTAAAGGGGTCCATCTGTCTACTGAAAAAGAGGTGCTGAGC 3986  
 DB 1510 GCTCTTACAGCACTGGGCCCCAGCAGACAGACAGGTCCTCTGAACCGAGACAGTTTGTCA 1569

QY 3987 TGCATCTTTTGCAGAGAAAGACAGAGGTGAAATAATAGAAAAATAATGCCATGTATATCG 4046  
 DB 1570 TGTATATTATGTCAGAGGAGCAAGAGGTGACTGTGGGAAGCAGGCGGATGCTTCTTGGCA 1629

QY 4047 GCCTGTGTCAGAAATCTACTGCTTAAACAGCAGGAGGAGAAACCCATAGAACTCTCA 4106  
 DB 1630 GCATTTGTCAGAGGTCAACCGTTCTGTCAAAGACAGGACGAAACCATTCGCGGACCCA 1689

QY 4107 GGAGAGCGCTTAGACCTTTTTCATGATCCAGACTTGGCATATGGAATTTATACAGCA 4166  
 DB 1690 GAAAA-----TATGATCCATTTATTCATGACCCCGATCTGTCTTGTGGGACACACATG 1746

QY 4167 AGCTGTGTCATGTAATGACGACAGTGTGCTGGCAGAGATTTTGTGAAGCTGTAC----- 4221







Kanagawa 230-0045, Japan (E-mail: genome-res@gsic.riken.jp,  
URL: http://genome.gsic.riken.jp/, Tel. 81-45-503-9222,  
Fax: 81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsic.riken.jp/  
URL: http://phantom.gsic.riken.jp/.

FEATURES  
Location/Qualifiers  
1..3058  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM DB:6430402B23"  
/db\_xref="taxon:10090"  
/clone="6430402B23"  
/sex="male"

CDS  
/tissue\_type="olfactory brain"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
-1..691  
/note="unnamed protein product; putative  
ubiquitin protein ligase E3 component n-recogin 1  
(MGDI|MGII:1277977, GB|NM\_009461, evidence: BLASTN, 100%,  
match=2560)"  
/codon\_start=2  
/protein\_id="BAC37160.1"  
/db\_xref="GI:26347023"  
/translation="GVAPPEELPANSAGEPSPALCSYLSPLNLFLLFOEYNDTIRPL  
LQWCGDPAKLSLQKSAVVRPRKNSLIPELSDYSLNQAHSFRCPSADDERK  
HPVLCIFCAILCSQICQIEVNGEVGACVPHALHCGAGVCIPLKRECRVLEVG  
KARGCAYPAPYLDEYGETDPLKRGNPLHLSRERYKLHVWQHCIIIEIARSQETN  
QMLFGFNWOLL"  
3032..3037  
/note="putative"  
3058  
/note="putative"

polyA\_signal

polyA\_site

ORIGIN

Query Match 11.7%; Score 740.2; DB 3; Length 3058;  
Best Local Similarity 82.4%; Pred. No. 8.9e-185;  
Matches 904; Conservative 0; Mismatches 148; Indels 45; Gaps 3;  
QY 5255 TGGGGTAACCTCCGCTGAGGAACCTGCATACCAATCTCGCAGAGGAGAGTACAGTGCAC 5314  
DB 1 TGGAGTAGCTCCGCTGAGNACTGTTGCCAATCTGCTGAGGAGAAATTCAGTGCAC 60  
QY 5315 CTGTAGCTATCTATCTTTACCTACCAATTTGTTCTGCTCTCCAGGAATATTTGGGATAC 5374  
DB 61 CTGTAGCTATCTATCTTTACCCACAAATTTGTTCTGCTCTTTCCAGGAATATTTGGGATAC 120  
QY 5375 TGTAAGCCCTCTCTCCAGAGGGTGTGCGAGATCTGCTTACTTAACATGTTTGAAGCA 5434  
DB 121 CATAAAGCCCTTACTACAGAGGGTGTGCGAGATCTGCTTACTCAAGTCTTTGAAGCA 180  
QY 5435 AAAAAACACCGTGTGAGGTACCTAGAAAAGAAATAGTTTGTATGAGCTTCTCTGATGA 5494  
DB 181 GAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240  
QY 5495 CTATAGCTGCTCTCTGATCAACATCTCTATTTTCCAGGTGCCCGCTGTGCGAGATGA 5554  
DB 241 CTACAGCTGTCTTTCTAAATCAGGCTTCTCCTTTAGGTGTCCAGGCTGTGCGAGATGA 300  
QY 5555 GCGAAGACCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 5614  
DB 301 GCGAAGACATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 360  
QY 5615 CTGCGAGAAATTTGTGAACCGGGAAGAGGTGTGAGCTTGTGATTTTTCACCACTTTCAC 5674  
DB 361 TTGCCAAGAAATAGTGAATCGGGAAGAGGTGTGAGCGGTGCTGTTTTCATCGCTTTCATTG 420



Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					</
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

evidence: FASTY, 78.8% ID, 100% length, match=749")  
/codon\_start=1  
/protein\_id="BAC38864.1"  
/db\_xref="GI:26350449"  
/translation="MFRYAVDILTWKESLPEDLEVAEKSDTYCYNLFNDEVHTYEQ  
VIYTLQAVNCTQKBAIPATVDRDRSRYGDFQYCDQAKTIVIRNTRSRQTKPLK  
VOMHSSVAHQNFGLKALSWGLSGVIGDGLRILCOVLOEGPDGENSSLDRLML  
NDSKLMKGSVYHQLPMSLLMDLYKXKLPAIRPAKNYERLQDYVTDHDEPFSVA  
DLSVQIFTPVPSLARMLBENLMTVIIKAFMDHLKRDQGRFQFERYTALQAFKFR  
VGLTILDLKVLISKEPWSDELRFQLOGFAPLELLKMQMGMDITRQVGHIME  
PEWEAAFTQIRYCVSOKSVIHLPIISRLLAGHLVLSKSEVAKYPPELLPLSLEPPL  
SICGHSEVETIRYCVSOKSVIHLPIISRLLAGHLVLSKSEVAKYPPELLPLSLEPPL  
LIHLPLRCLVCAQVHWRNGFSLVNOIYYVHNKCRREMFDKDIYMLQTVGSMW  
DNHFLMILSRPELQLPSTPDYGRKPSSESVTHKDVQVQNNLILBEMLYLILVGE  
RNFPGVQVADITIKREIHLQSLIKPMFSELSVLPSIDENKETMESVIESVAHFKE  
KPLCTGRMVLPEKCAKGFNLYFYHRAEQSKABEQKLRKREKBDTALPPPALP  
PFCLFASLIVILQCDMLYIMGTILQWAVEHHSQSAWSEMLQRLVHLIEMALQEBKH  
HLENAVEGHVOTFTFKIKSPGDAPHNSPSILAMLETIQNAPSLBAHKMDIRWLKLM  
FNAIKKIRECSSSPVAEAEAGTMESSRDKKAEKRKRAEIAIRLREKIMAQMPFMO  
RHPD"

## ORIGIN

Query Match 10.9%; Score 687.8; DB 3; Length 3627;  
Best Local Similarity 56.0%; Pred. No. 8.4e-171;  
Matches 1472; Conservative 0; Mismatches 1107; Indels 51; Gaps 7;  
QY 1245 GAAGAGTAAATGTCAGGCGAGAAATAATTCCTTCAGTGATAAATAATGTCGTAGAA 1304  
DB 1005 GAAGATGTATGCCAGAACTTACAACTTTTGTATTATGTTTCGATACAGTAGAT 1064  
QY 1305 ATGACTATATGGAAGGAAAGAACTCCCTCGTGAATCCAGATAGGGAGAAAT 1364  
DB 1065 ATACTGACCTGGGAAAGAAAGTGAATTCCTGGAAGACTTAGAAGTGGCAGAGAGAT 1124  
QY 1365 GAAAGATATATATGTCCTTTCAATGATGAACACCATTCATATGACACAGCTATATAC 1424  
DB 1125 GACACCTACTGATGCTGTTTATGATGAGGTTACACCTATGAGCAGATCATTTAT 1184  
QY 1425 AGCCTCAAAAGACTCTTGATGAGCTGCGAGAGCCAGTTCGATACCACTGCCATT 1484  
DB 1185 ACCCTTCAGAAAGCTGTGAACGTACACAGAGGAAGCCATTCGCTTGGCACTACAGTT 1244  
QY 1485 GACAAAGAGGTCGTGGGCTGTTAAGCGGAGCTTATGCTGCTGCCAGGAGCAAG 1544  
DB 1245 GATCGAGATGCGCGTAGGTCGTCCGATATGGAAGATTTCCAGTACTGTGATCAAGCAAG 1304  
QY 1545 GAAGATATAAGAGTCAATCAGAAATGTCTCTCAACATCCATTCATGTAGAGTATTA 1604  
DB 1305 ACAGTCAATGTGAGGAACACCGACAGACAGACC---AGCGCTCAAGTTCAAGTTATG 1361  
QY 1605 CACTCAGAGATTATGCTCATCAGAAATTTGCTTTGCGTCTGCTGCTGGCTGGGAAGT 1664  
DB 1362 CACTCCTCGTGGCTGCTCATCAGAAATTTGCTTTGAAAGCTCTGCTGCTGGCTGGGAAGT 1421  
QY 1665 ATTATGAGTATTCAGTGAATTTAGGAGCATCTTTGGCAGCATGCTTCAGAGAGAA 1724  
DB 1422 GTTATTTGGATCTCAGATGCGCTTCAGAGATTTTGTCTCAAGTTTGGATTACAGAGAGT 1481  
QY 1725 CCTGACTCGGAGAAATCCCTGTCTCATAAGCAGGTAAATGCTTTGGGATGCAAGCTTTAT 1784  
DB 1482 CCAGATGGGGAAGAACTCTCTCTGCTGGTGAACAGACTGATGCTTAATGATTCGAATATGG 1541  
QY 1785 AAAGGTGCCGTAAGATCTCTCATGAAATTTGATCTTCAGCAGATTTTTTTTATGGAGATGGA 1844  
DB 1542 AAAGGGCTAGGAGTGTGATACACAGTTGTTTCATGAGCAGCTGCTCATGGACCTCAAG 1601  
QY 1845 TACAAAAAATCTTTTGTATGGAATTTGGAAGTATTTATAACAACTGCGAGAAAGATAT 1904  
DB 1602 TATAAGAGCTGTTCGCGCTTCGATTTGCTTAAAGAACTATGAGCGTTTTCAGAGAGATAT 1661  
QY 1905 ATCAGTGTATCATCAGCAGAGTATCTCTATAACTGCACTTCAGTTTCAGTGTGTTTACT 1964  
DB 1662 GTGACAGATGACCAACAGAGAGATTTTTCAGTGGCAGACCTCTCGGTTTCAGATATTCACA 1721

QY 1965 GTTCTCTACTCTGCTCGACATCTTATTGAGAGCAGAAATGTTATCTCTGTCTATTACTGAA 2024  
DB 1722 GTTCTCTCTACTCTGCTCGAAATGCTCTCTCACAGAGAGAACCTGATGACCGTTATCATTAAG 1781  
QY 2025 ACTCTGCTAGAAAGTTTACTCT-----GAGTACTCTGACAGAGAAACAATAAATCAACTTC 2078  
DB 1782 GCTTTCATGACCAATTTGAAACACAGAGATGCCAGGGCAGATTCGATTTGAGACGCTAC 1841  
QY 2079 CAGGGTTATAGCCAGACAAAATTTGGGAAGATATATGCAAGTAAATATATGACCTTAAAGTAT 2138  
DB 1842 ACTGCCCTCCAAAGCTTCAAGTTTCAGGAGAGTCCAGAGCCTCATCTTAGATCTCAAGTAT 1901  
QY 2139 ATCTGATCAGCAAAACCCACAATATGACAGAAAGATTTAAGATGCAAGTTCCTTTGAAGGT 2198  
DB 1902 GTATTGATTTAGCAAAACCAACGAGAGTGTGATGAGCTGAGGCGAAGATTTCTTACAAGGG 1961  
QY 2199 TTTGAGATCTTTTGAAGATTTCTTACTCTGATGCGGGAATGGAAGAAATCCGAAAGCAG 2258  
DB 1962 TTGATGCTCTTCTTGGAAATTTACGAAGTGCATGCGGGAATGGAACCGGATCAGCGCTCAG 2021  
QY 2259 GTTGGCAAACATTTGAACTGGATCTGATTTGGAGGCTGCCATTCCTATACAGATGCAA 2318  
DB 2022 GTGGGACAGACATTTGAGATGAGCCAGAGTGGGAAGCAGCTTTCACACTGACAGATGAAG 2081  
QY 2319 TTGAAGAAATATTTTACTCATGTTCCAGAGTGTGCTGTTGATGAGAACTCTTACTTT 2378  
DB 2082 CTGACACAGCTCATCTCAATGTTGTCAGGACTGCTGTGCTCTGGACGAAAAAGTGTAAAT 2141  
QY 2379 GTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGGTGAGTACAGTTCATATCTAGT 2438  
DB 2142 GAAGCTTACAGAAATATGCTGCTGCTGACACAGTGTCAAGCGGGAATTTACTGATGTT 2201  
QY 2439 AGCAAGACAGT---AGTACAATCTGTGTGACATATGTTTGGAAACAAAGTCTTACAGAGTA 2495  
DB 2202 GAACAGCCAATCACACTCAGTATTTGTGACACTCGGTGGAAACCATCAGATCTGTGTT 2261  
QY 2496 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTTAGAGACCTTCTGCTGCTTTCATGTA 2555  
DB 2262 TCCCAAGAAAAAGTTAGCAATTCACCTCCCAATTTCTGCTTGTCTGCTGAGGTTTGCATGTA 2321  
QY 2556 CGTTTAAAGCAGGCTGGTCTCTTCAAGACTGCATGAATTTGCTGCTTTTGGAGGACTTT 2615  
DB 2322 TTGTTAAGAAAAAGTGAAGTGGCATATAAATTTCCAGAGCTCCTACTCTTAAGTGAACATG 2381  
QY 2616 CAAGTAGAGTACTAGTGAATATCTTTACGTTGTCTGCTGTTGCTGCTGCCAGGTTGTT 2675  
DB 2382 AGCCCAACCATGTTGATAGAACTCTCTTAGATGCTTGTCTTATGCTCAAGTGCAT 2441  
QY 2676 GCTGAGATGTGCGAAGAAATGGAATCTCTTATTTAGCCAGGTGTTTATTTACCAAGAT 2735  
DB 2442 GCTGGGATGTGGAAGAAATGGCTTCTCTTAGTAAATCAGATCTATTACTACCAATAAT 2501  
QY 2736 GTTAAAGTGCAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2795  
DB 2502 GTGAAATGCGAGGAGAGATGTTTCGACAGGACATAGTATGCTTTCAGACAGGTGCTCC 2561  
QY 2796 TTAATGGATCCCAATTAAGTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2855  
DB 2562 ATGATGACCCCAACCACTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2621  
QY 2856 TTTAAACAGACCATATCTACAAA-----AGACAGGATTTGATTT 2894  
DB 2622 TTCAGCAGCCTGATGATGGAAGAGATTCAGTTCCTGAGGTTTACCCATAGGACGCTGTT 2681  
QY 2895 AAAAATATATACATATAGAGAAATGCTTTCAGGCTCTCATCTATATTTGGGTGAG 2954  
DB 2682 CAGCAGAAACACACTCTGATCGAAGAGATGCTCTACCTCATCATCTGCTTGTGGGAGAA 2741  
QY 2955 CGTTATGTACCTGGAGTGGAAATGTGACAAAAGAGAGGTCAATATGAGAGAAATCAAT 3014  
DB 2742 AGATTCAACCTTGGGTTGGACAGGTTGGCTGCCAAGATGAATTCAGAGGGGAGATATTC 2801





Db 361 GGAAATGTGAACGGGAGAGGTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGGAGC 420  
Qy 5681 CGGAGTCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGGTAAGC 5740  
Db 421 CGGAGTCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGGTAAGC 480  
Qy 5741 CAGAGGCTGCTATCCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTGGCT 5800  
Db 481 CAGAGGCTGCTATCCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTGGCT 540  
Qy 5801 GAAGAGGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGAAGCTCCATTTGGTCTG 5860  
Db 541 GAAGAGGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGAAGCTCCATTTGGTCTG 600  
Qy 5861 GCAACAACACTGCTATTAAGAAGATTCGTAGAGCCCAAGAGACTAATCAGATGTATT 5920  
Db 601 GCAACAACACTGCTATTAAGAAGATTCGTAGAGCCCAAGAGACTAATCAGATGTATT 660  
Qy 5921 TGGATTCAGTGGCAGTACTG 5942  
Db 661 GGGATTCACCCCTGCAGTACTG 682

## RESULT 13

BU155953  
LOCUS  
DEFINITION AGENCOURT\_7966363 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6165232 898 bp mRNA linear EST 03-SEP-2002 5', mRNA sequence.

ACCESSION BU155953

VERSION BU155953.1 GI:22669485

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC/DCTD/BTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LIAW13523 row: j column: 17

High quality sequence stop: 606.

Location/Qualifiers

1. .898

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6165232"

/tissue\_type="melanotic melanoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC-72"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;

Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dr.

Average insert size 2 kb. Library constructed by Life

Technologies."

ORIGIN

Query Match 10.6%; Score 668.8; DB 5; Length 898;

Best Local Similarity 99.6%; Pred. No. 5.9e-166;

Matches 681; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 5626 TTGTGAACGGGAGAGGTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGGAGCGGAG 5685

Db 12 TTGTGAACGGGAGAGGTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGGAGCGGAG 71

Qy 5686 TCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGGTAAGCAGAG 5745  
Db 72 TCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGGTAAGCAGAG 131  
Qy 5746 GCTGTGCCCTATCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTGGCTCAAGA 5805  
Db 132 GCTGTGCCCTATCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTGGCTCAAGA 191  
Qy 5806 GGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGGAAGCTCCATTTGGTCTGCAAC 5865  
Db 192 GGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGGAAGCTCCATTTGGTCTGCAAC 251  
Qy 5866 AACACTGCATTTAGAGAGATTCGTAGAGCCCAAGAGACTAATCAGATGTATTGGAT 5925  
Db 252 AACACTGCATTTAGAGAGATTCGTAGAGCCCAAGAGACTAATCAGATGTATTGGAT 311  
Qy 5926 TCAACTGGCAGTACTGTGAGCTCCAACTCTGAGCTCCAACTCTGAGCGCTGGGAAAGTATTGGAAGGT 5985  
Db 312 TCAACTGGCAGTACTGTGAGCTCCAACTCTGAGCTCCAACTCTGAGCGCTGGGAAAGTATTGGAAGGT 371  
Qy 5986 AGTAAAGGCTGATTCAAATTTATGAAATCTTCTGAGCGCTGGGAAAGTATTGGAAGGT 6045  
Db 372 AGTAAAGGCTGATTCAAATTTATGAAATCTTCTGAGCGCTGGGAAAGTATTGGAAGGT 431  
Qy 6046 CTTTGTCTCCATGTCAGGTTCACTTACATCAATTAATATTTCTTAATGGAGTATTGCT 6105  
Db 432 CTTTGTCTCCATGTCAGGTTCACTTACATCAATTAATATTTCTTAATGGAGTATTGCT 491  
Qy 6106 TTCAATTAGCAAAACATATGCTTTCAGAGGAAAGGACATAGATCAATCTGTTTTATG 6165  
Db 492 TTCAATTAGCAAAACATATGCTTTCAGAGGAAAGGACATAGATCAATCTGTTTTATG 551  
Qy 6166 CTAGTATTTCCAGGAATTTTATTCCTTCATATTTTGTCTCATTTTCAATTTTTCATCC 6225  
Db 552 CTAGTATTTCCAGGAATTTTATTCCTTCATATTTTGTCTCATTTTCAATTTTTCATCC 611  
Qy 6226 ACTGCTAGATGAGTCAGTCAGTCAAAAGTGTAGACATTTTATGTTT-GGTTAACTCTT 6284  
Db 612 ACTGCTAGATGAGTCAGTCAGTCAAAAGTGTAGACATTTTATGTTTGGTGGTAACTCTT 671  
Qy 6285 CTGCAATTTTGTATTGGTGTGTTT 6308  
Db 672 CTGCCATTTTGTATTGGGTTTT 695

## RESULT 14

BU18516/c

LOCUS

DEFINITION

UI-H-PH1-bfk-c-21-0-UI.s1 NCI CGAP FHI Homo sapiens cDNA clone

UI-H-PH1-bfk-c-21-0-UI 3', mRNA sequence.

ACCESSION

BU18516

VERSION

BU18516.1

KEYWORDS

EST.

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 668)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: James Martin

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

FEATURES

```
source
1. .668
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FH1-bfk-c-21-0-UI"
/tissue_type="Cell Line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_FH1"
/note="Organ: Chondrosarcoma; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FH1 is a normalized cDNA library obtained from a cell line derived from grade I chondrosarcoma tissue. The library was constructed and normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCCGCG. The cell line was provided by Dr. James Martin from the University of Iowa.
TAG_RISSE=Human Chondrosarcoma Cell Line CS8 - Grade I
Chondrosarcoma
TAG_LIB=UI-H-FH1
TAG_SEQ=AGAATCCGCG"

ORIGIN
Query Match 10.5%; Score 664.8; DB 5; Length 668;
Best Local Similarity 99.9%; Pred. No. 6.2e-165;
Matches 666; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2607 GAGGACTTTCACTAGAGGTTAGTGGAAATCTTTTACGTTGTCTGGTGTGGTTGCC 2666
DB 668 GAGGACTTTCACTAGAGGTTAGTGGAAATCTTTTACGTTGTCTGGTGTGGTTGCC 609
QY 2667 CAGGTTGTCTGAGATGGCGAAGAAATGGATCTCTCTTATTAGCCAGGTTGTTTAT 2726
DB 608 CAGGTTGTCTGAGATGGCGAAGAAATGGATCTCTCTTATTAGCCAGGTTGTTTAT 549
QY 2727 TACCAGATGTTTAACTGCAGAGAGAAATGTATGATAAGATATCATCTTCAGATT 2786
DB 548 TACCAGATGTTTAACTGCAGAGAGAAATGTATGATAAGATATCATCTTCAGATT 489
QY 2787 GGTGCATCTTTAATGATCCCAATAGTTCTTCTTACTGTACTTTCAGAGGTATGAATT 2846
DB 488 GGTGCATCTTTAATGATCCCAATAGTTCTTCTTACTGTACTTTCAGAGGTATGAATT 429
QY 2847 GCCGAGGCTTTTAAAGACCATATCTACAAAGACAGGATTTGATTTAAACAATATAT 2906
DB 428 GCCGAGGCTTTTAAAGACCATATCTACAAAGACAGGATTTGATTTAAACAATATAT 369
QY 2907 ACATTAATGAAGAAATGTTTCAAGTCTCTCATCTATATTGTTGGTGCAGGTTATGACCT 2966
DB 368 ACATTAATGAAGAAATGTTTCAAGTCTCTCATCTATATTGTTGGTGCAGGTTATGACCT 309
QY 2967 GGAGTGGGAAATGTGACCAAGAGAGGTTCACATGAGAGAAATCAATCACTTGTCTTTC 3026
DB 308 GGAGTGGGAAATGTGACCAAGAGAGGTTCACATGAGAGAAATCAATCACTTGTCTTTC 249
QY 3027 ATTGAACCCATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGATGAAAATATATGA 3086
DB 248 ATTGAACCCATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGATGAAAATATATGA 189
QY 3087 ACTGGCTTACAGATGTCATAAACAAGTGGCCACATTTAAGAAACAGGTGATCAGGC 3146
DB 188 ACTGGCTTACAGATGTCATAAACAAGTGGCCACATTTAAGAAACAGGTGATCAGGC 129
QY 3147 CATGGAGTTTATGAACTTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 3206
DB 121 CATGGAGTTTATGAACTTAAAGATGAATCACTGAAAGAGGTTCATATGAGAGAAATCATTC 180

Db 128 CATGGAGTTTATGAACTTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 69
QY 3207 TACTCCAAACCCAGCATAGCAAGCTCAACATATGCAAGAGAGAGAGAGAGAGAGAG 3266
DB 68 TACTCCAAACCCAGCATAGCAAGCTCAACATATGCAAGAGAGAGAGAGAGAGAGAG 9
QY 3267 AACAAAGA 3274
DB 8 AACAAAAA 1

BG534574 756 bp mRNA linear EST 03-APR-2001
602553425F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4663182 5',
mRNA sequence.
BG534574
ACCESSION BG534574
VERSION BG534574.1 GI:13526116
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1465 row: a column: 07
High quality sequence stop: 751.
Location/Qualifiers
1. .756
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4663182"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccgctcgcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTTATGCG-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCGGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 10.4%; Score 656.6; DB 4; Length 756;
Best Local Similarity 95.9%; Pred. No. 9.9e-163;
Matches 730; Conservative 0; Mismatches 19; Indels 12; Gaps 5;

QY 2836 GGTATGAACTTCCCGAGGCTTTTAAACAGACCATATCTACAAAGACAGGATTTGATT 2895
DB 1 GGTATGAACTTCCCGAGGCTTTTAAACAGACCATATCTACAAAGACAGGATTTGATT 60
QY 2896 AACAAATAATACACTAATAGAGAAATGCTTCAGTCTCTCATCTATATTGTTGGTGC 2955
DB 61 AACAAATAATACACTAATAGAGAAATGCTTCAGTCTCTCATCTATATTGTTGGTGC 120
QY 2956 GTTATGTACCTGGAGTGGGAAATGTGACCAAGAGAGAGGTTCACATGAGAGAAATCATTC 3015
DB 121 GTTATGTACCTGGAGTGGGAAATGTGACCAAGAGAGAGGTTCACATGAGAGAAATCATTC 180
```

```
QY 3016 ACTTGCTTTGCATTGAACCCATGCCACAGTGGCCATTGCCAAAAATTTACCTGAGATG 3075
Db |||||||
QY 181 ACTTGCTTTGCAATTGAACCCATGCCACAGTGGCCATTGCCAAAAATTTACCTGAGATG 240
Db |||||||
QY 3076 AAAATAATGAAACTGGCTTAGAGAAATGTCATAAACAAGTGCCACATTTAAGAAACCCAG 3135
Db |||||||
QY 241 AAAATAATGAAACTGGCTTAGAGAAATGTCATAAACAAGTGCCACATTTAAGAAACCCAG 300
Db |||||||
QY 3136 GTGTATCAGGCCATGGAGTTTATGAACATAAAGATGAATCACTGAAAGACTTCAATATGT 3195
Db |||||||
QY 301 GTGTATCAGGCCATGGAGTTTATGAACATAAAGATGAATCACTGAAAGACTTCAATATGT 360
Db |||||||
QY 3196 ACTTTTATCAATTAATCCAAACCCAGCATAGCAAGCTGAACATATGCGAAGAAAGGA 3255
Db |||||||
QY 361 ACTTTTATCAATTAATCCAAACCCAGCATAGCAAGCTGAACATATGCGAAGAAAGGA 420
Db |||||||
QY 3256 GAAAAACAAGAAAAACAAGATGAAG-----CATTTGCGGCCACCACCCTCTGAAATTC 3309
Db |||||||
QY 421 GAAAAACAAGAAAAACAAGATGAAGTTAAAAACAATGCGCGCCACCACCCTCTGAAATTC 480
Db |||||||
QY 3310 GCCCTGCTTTTCAGCAAGTGATTAACCTTCTCAACTGTGATAT-CATGATGTACATTC 3368
Db |||||||
QY 481 GCCCTGCTTTTCAGCAAGTGATTAACCTTCTCAACTGTGATATCCATGATGTACATTC 540
Db |||||||
QY 3369 AGGACCGTATTTGAGCGGGCAATAGACACAGATTCCTAACTTGTGACCGAAGGGATGCTC 3428
Db |||||||
QY 541 AGGACCGTATTTGAGCGGGCAATAGACACAGATTCCTAACTTGTGACCGAAGGGATGCTC 600
Db |||||||
QY 3429 CAAATGGCTTTTCATATTTCTGGCATTTGGGTTTACTAGAGAGCAAGCAAGCTTCAAAA 3488
Db |||||||
QY 601 CAAATGGCTTTTCATATTTCTGGCATTTGGGTTTACTAGAGAGCAAGCAAGCTTCAAAA 659
Db |||||||
QY 3489 GCTCCTGAAGAGCAAGTAACATTTGACTTTTATCATAAGGCTTCAAGATTTGGAGTTCA 3548
Db |||||||
QY 660 GCTCCTGAAGAGCAAGTAACATTTGACTTTTATCATAAGGCTTCAAGATTTGGAGTTCA 719
Db |||||||
QY 3549 GCCATGAATATACAAATGCTTTTGGAAAAAATCAAGGAAT 3589
Db |||||||
QY 720 G-CATGAATATACAA---TGTTTTGGAAAAAATCAAGGAT 756
Db |||||||
```

Search completed: July 9, 2005, 07:46:27  
Job time : 17604 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 07:46:33 ; Search time 130 Seconds  
(without alignments)  
5203.415 Million cell updates/sec

Title: US-10-758-672A-2  
Perfect score: 9224  
Sequence: 1 MADEAGGTERMEISAEPLQ.....ETARQETNQLFGFNWQLL 1749

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	ADJ95442	Adj95442 Human Ubi
2	9224	100.0	1749	ADJ95442	ADJ95442 Human Ubi
3	9168	99.4	1738	ADJ95442	ADJ95442 Human Ubi
4	9087	98.5	1734	ADJ95459	ADJ95459 Human Ubi
5	9087	98.5	1734	ADJ95459	ADJ95459 Human Ubi
6	8617	93.4	1757	AAW84351	AAW84351 Murine ub
7	8617	93.4	1757	AAW84351	AAW84351 Murine ub
8	8617	93.4	1757	AAW84351	AAW84351 Murine ub
9	8617	93.4	1757	AAW84351	AAW84351 Murine ub
10	4768	51.7	908	ADJ95442	ADJ95442 Human Ubi
11	4327	46.9	1755	ADJ95442	ADJ95442 Human Ubi
12	4327	46.9	1755	ADJ95442	ADJ95442 Human Ubi
13	4278	46.4	1755	ADJ95446	ADJ95446 Mouse Ubi
14	4209	45.6	811	ADJ95446	ADJ95446 Mouse Ubi
15	4209	45.6	811	ADJ95446	ADJ95446 Mouse Ubi
16	3457	37.5	1400	AAW78576	AAW78576 Human pro
17	3454	37.4	1400	AAW78576	AAW78576 Human pro
18	3160	34.3	1275	ADJ95442	ADJ95442 Human Ubi
19	3066	33.2	594	ADJ95497	ADJ95497 Human pro
20	2710.5	29.4	1109	AAW02376	AAW02376 Polypepti
21	2278.5	24.7	1829	ADJ95442	ADJ95442 Human Ubi
22	2168	23.5	424	ADJ95442	ADJ95442 Human Ubi
23	1716	18.6	333	AAW84353	AAW84353 Partial h
24	1716	18.6	333	AAW84353	AAW84353 Partial h
25	1670	18.1	306	AAW81609	AAW81609 Tumour-as

26	1665	18.1	306	8	ADR09273	Human pro
27	1301.5	14.1	575	8	ADJ95442	Novel hum
28	1169.5	12.7	452	5	ADJ95442	Human pol
29	903.5	9.8	343	8	ADJ95442	Human pro
30	877	9.5	487	4	ADJ95442	Novel hum
31	790	8.6	258	4	ADJ95442	Novel hum
32	606.5	6.6	1225	5	ADJ95442	Herbicida
33	468	5.1	147	7	ADJ95442	Human pro
34	453.5	4.9	741	8	ADJ95442	Human pro
35	453.5	4.9	951	6	ADJ95442	Human pro
36	449	4.9	389	4	ADJ95442	Novel hum
37	420	4.6	1456	4	ADJ95442	Drosophil
38	414.5	4.5	203	5	ADJ95442	Human pol
39	404.5	4.4	709	8	ADJ95442	Human pro
40	401	4.3	71	8	ADJ95442	Mouse N-r
41	360	3.9	153	4	ADJ95442	Human imm
42	341	3.7	550	8	ADJ95442	Human pro
43	311	3.4	783	4	ADJ95442	Drosophil
44	311	3.4	783	8	ADJ95442	Drosophil
45	302	3.3	437	8	ADJ95442	Novel hum

## ALIGNMENTS

## RESULT 1

ADJ95442  
ID ADJ95442 standard; protein; 1749 AA.

AC ADJ95442;

DT 03-JUN-2004 (first entry)

DE Human Ubiquitin ligase E3alpha I.

KW Human; enzyme; ubiquitin ligase; E3alpha I; ubiquitin-proteasome pathway;  
KW Gene therapy; vaccine; muscular atrophy; cachexia; catabolic disorders;  
KW cancer cachexia; renal cachexia; inflammatory cachexia;  
KW muscle wasting disorder; metabolic acidosis; uremia; burn;  
KW hyperthyroidism; Cushing's syndrome; fasting; denervation atrophy;  
KW diabetes mellitus; sepsis; AIDS wasting syndrome; SNP;  
KW single nucleotide polymorphism.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 1568  
/note= "May be Arg as the result of a single nucleotide polymorphism"

US6706505-B1.

16-MAR-2004.

28-NOV-2000; 2000US-00724126.

08-MAR-2000; 2000US-0187911P.

(AMGE-) AMGEN INC.

Han H, Kwak K;

WPI; 2004-236723/22.

N-PSDB; ADJ95441.

New nucleic acid molecule, useful for preparing a composition for  
diagnosing, treating or preventing diseases associated with human  
E3alpha polypeptide, e.g., muscle atrophy.

Claim 1; SEQ ID NO 2; 104pp; English.

The invention relates to a new isolated nucleic acid molecule appearing  
as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I



CC protein appearing as ADJ95442. Also included are a vector comprising the  
 CC nucleic acid, a host cell comprising the vector, a process of producing a  
 CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the  
 CC nucleic acid molecule, a reagent comprising a detectably labelled  
 CC nucleotide, and a method for determining the presence of a human E3alpha  
 CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid  
 CC molecule is useful for preparing a composition for diagnosing, treating  
 CC or preventing diseases associated with human E3alpha I polypeptide, e.g.  
 CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal  
 CC cachexia, inflammatory cachexia, muscle wasting disorders associated with  
 CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,  
 CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting  
 CC syndrome. The present sequence represents human E3alpha I.  
 XX  
 SQ

Query Match	100.0%;	Score 9224;	DB 8;	Length 1749;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1749;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MADDEAGGTERMEISAELEPQTPORLASWDDQVDFYTAFLHLAQLVPEIYFAEMDPDLE	60		
DB 1	MADDEAGGTERMEISAELEPQTPORLASWDDQVDFYTAFLHLAQLVPEIYFAEMDPDLE	60		
QY 61	KOBESVQMSIPTLEWYLFGEEDPDI CLEKLKHSQAFOLGCRVFKSGETTYSRDCALDPT	120		
DB 61	KOBESVQMSIPTLEWYLFGEEDPDI CLEKLKHSQAFOLGCRVFKSGETTYSRDCALDPT	120		
QY 121	CVLCMDCFDQDSVHKHRYKMHNTSTGGGFCDCGTEAWKTGPFVKNHPEGRAGTIKENSRC	180		
DB 121	CVLCMDCFDQDSVHKHRYKMHNTSTGGGFCDCGTEAWKTGPFVKNHPEGRAGTIKENSRC	180		
QY 181	PLNEEVIQARKIPPSVIKVVEMTWEBEKELPPELOIREKNERYCYLVFNDHSHYDH	240		
DB 181	PLNEEVIQARKIPPSVIKVVEMTWEBEKELPPELOIREKNERYCYLVFNDHSHYDH	240		
QY 241	VYISLQALDCELAELAAQLHTTADKEGRRAVAGAAQAEAKEDIKSHSENVSHPLHV	300		
DB 241	VYISLQALDCELAELAAQLHTTADKEGRRAVAGAAQAEAKEDIKSHSENVSHPLHV	300		
QY 301	EVLHSEIMAHQKFPALRGSMNKMINSYSSDFRQIFCOACLRBEPDSENCLISRLMLMDA	360		
DB 301	EVLHSEIMAHQKFPALRGSMNKMINSYSSDFRQIFCOACLRBEPDSENCLISRLMLMDA	360		
QY 361	KLYKGARKILHELI FSSPFMEYKCLPAMEFVKYKQLOKEVISDDHRSISITALSVO	420		
DB 361	KLYKGARKILHELI FSSPFMEYKCLPAMEFVKYKQLOKEVISDDHRSISITALSVO	420		
QY 421	MFTVPTLARHLIEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK	480		
DB 421	MFTVPTLARHLIEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK	480		
QY 481	YILISKPTIWTERRMQFLGFRSFLKILTCMQMBEIRRQVGOHIEVDPDWEAAIAIQM	540		
DB 481	YILISKPTIWTERRMQFLGFRSFLKILTCMQMBEIRRQVGOHIEVDPDWEAAIAIQM	540		
QY 541	QLKNILLMFQEWACDBELLVAYKECHKAVMRCSTSFSSKTVVQSCHSLETYSYV	600		
DB 541	QLKNILLMFQEWACDBELLVAYKECHKAVMRCSTSFSSKTVVQSCHSLETYSYV	600		
QY 601	SEDLVSIHLPLSTLGLVRLSRGAVSRHLHFVSFEDQVRLVEYPLRCLVLVAQVY	660		
DB 601	SEDLVSIHLPLSTLGLVRLSRGAVSRHLHFVSFEDQVRLVEYPLRCLVLVAQVY	660		
QY 661	AEWRRNGLSLISQVFFYQDVCKREEMDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEA	720		
DB 661	AEWRRNGLSLISQVFFYQDVCKREEMDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEA	720		
QY 721	FNKTIKTDQDLIKQNTLIEEMQLVLIYIVGSRYPGVGNVTKEEVTWREIIHLLCIEP	780		
DB 721	FNKTIKTDQDLIKQNTLIEEMQLVLIYIVGSRYPGVGNVTKEEVTWREIIHLLCIEP	780		
QY 781	MPHSAIAKNLPENENNETGLENVINKVATPKPGVSGHGVYELKDESLKDFNMYFYHYSK	840		

DB 781	MPHSAIAKNLPENENNETGLENVINKVATPKPGVSGHGVYELKDESLKDFNMYFYHYSK	840		
QY 841	TOHSAEAMOKKRRQENKDEALPPPPPEPCPAFSKVINLLNCDDIMMYILRTVPERAID	900		
DB 841	TOHSAEAMOKKRRQENKDEALPPPPPEPCPAFSKVINLLNCDDIMMYILRTVPERAID	900		
QY 901	TDSNLTWTEGMLQMAFHIALGLLEBEKQOLQKAPBEVTFDFVHKASRLGSSAMNQLMLE	960		
DB 901	TDSNLTWTEGMLQMAFHIALGLLEBEKQOLQKAPBEVTFDFVHKASRLGSSAMNQLMLE	960		
QY 961	KLKGIPLQEGQKDMITWILQMPDTPVKRLREKSKLIVATTSGSESINKDEIHTHDKBAERK	1020		
DB 961	KLKGIPLQEGQKDMITWILQMPDTPVKRLREKSKLIVATTSGSESINKDEIHTHDKBAERK	1020		
QY 1021	RKAEAAARLHROKIMAAQMSALQKNFIETHKLMYDNTSEMPGKEDSIMESESTPAVSDYSRI	1080		
DB 1021	RKAEAAARLHROKIMAAQMSALQKNFIETHKLMYDNTSEMPGKEDSIMESESTPAVSDYSRI	1080		
QY 1081	ALGPXRGPSVTEKEVLTCTILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA	1140		
DB 1081	ALGPXRGPSVTEKEVLTCTILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA	1140		
QY 1141	LDPLFMDPDLAYGTYTSGCGHVMHACWQKYFEAVQLSSQQRHIVDLFDLESSEYLCPLC	1200		
DB 1141	LDPLFMDPDLAYGTYTSGCGHVMHACWQKYFEAVQLSSQQRHIVDLFDLESSEYLCPLC	1200		
QY 1201	KSLCNTVPIPIPLQPKINSENADALAQLLTARMIQTVLARISYNTIRHAKGENPIPIF	1260		
DB 1201	KSLCNTVPIPIPLQPKINSENADALAQLLTARMIQTVLARISYNTIRHAKGENPIPIF	1260		
QY 1261	FNQMGDSTLRFHSITLSFGVSESSIKYSNSIKEMWTLFATTIYRIGLKVPPDPRDPRVPM	1320		
DB 1261	FNQMGDSTLRFHSITLSFGVSESSIKYSNSIKEMWTLFATTIYRIGLKVPPDPRDPRVPM	1320		
QY 1321	TWSTCAFTIQAIENTLLGDEGKPLFGALQNRQHNGLKALMQPAVAQRITCPOVLIOKHLVR	1380		
DB 1321	TWSTCAFTIQAIENTLLGDEGKPLFGALQNRQHNGLKALMQPAVAQRITCPOVLIOKHLVR	1380		
QY 1381	LLSVVLPMIKSEDTCLLSIDLPHVLVGAVALPSPSYWDDPDVLOPSSVSSSYNHLVLP	1440		
DB 1381	LLSVVLPMIKSEDTCLLSIDLPHVLVGAVALPSPSYWDDPDVLOPSSVSSSYNHLVLP	1440		
QY 1441	LITMAHMLQIILLTVDTGLPLAQVQDSSEAHSSASSFFAEISQYTSIGCDIPGWYLVWS	1500		
DB 1441	LITMAHMLQIILLTVDTGLPLAQVQDSSEAHSSASSFFAEISQYTSIGCDIPGWYLVWS	1500		
QY 1501	LKNGITPYLRCALPFHYLLGVTTPPEELHTNSAEGEYSALCSYLSLPTNLFLLPQBYWDT	1560		
DB 1501	LKNGITPYLRCALPFHYLLGVTTPPEELHTNSAEGEYSALCSYLSLPTNLFLLPQBYWDT	1560		
QY 1561	VRPILQRCADPALLNCLKQNTVYRPRKNSLIELPDYSCLLNQASHFCRPSADDE	1620		
DB 1561	VRPILQRCADPALLNCLKQNTVYRPRKNSLIELPDYSCLLNQASHFCRPSADDE	1620		
QY 1621	RKHPVLCIFCGAILCSQNICCOEIVNGBEVGCIFLHLCGAGVCIFLKIRECRVVLVSG	1680		
DB 1621	RKHPVLCIFCGAILCSQNICCOEIVNGBEVGCIFLHLCGAGVCIFLKIRECRVVLVSG	1680		
QY 1681	KARGCAYPAPYLDEYGETDPGLKRGNPPLHLSRERYKHLVWQOHCIIIEIARSQETNQ	1740		
DB 1681	KARGCAYPAPYLDEYGETDPGLKRGNPPLHLSRERYKHLVWQOHCIIIEIARSQETNQ	1740		
QY 1741	LFQFNWQLL 1749			
DB 1741	LFQFNWQLL 1749			

RESULT 2  
 ID ADS86864 standard; protein; 1749 AA.  
 XX  
 AC ADS86864;

XX DT 16-DEC-2004 (first entry)  
 XX DE Human E3alpha ubiquitin ligase, hE3alpha1 protein sequence.  
 XX KW Human; E3alpha ubiquitin ligase; huE3alpha1;  
 KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;  
 KW fasting; metabolic acidosis; muscle degeneration; kidney failure;  
 KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;  
 KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;  
 KW inflammatory cachexia; hyperthyroidism; denervation atrophy;  
 KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;  
 KW gene therapy; enzyme.  
 XX OS Homo sapiens.  
 XX PN US2004185037-A1.  
 XX PD 23-SEP-2004.  
 XX PF 15-JAN-2004; 2004US-00758672.  
 XX PR 08-MAR-2000; 2000US-0187911P.  
 XX PR 28-NOV-2000; 2000US-00724126.  
 XX PA (HANH/) HAN H.  
 XX PA (KWAK/) KWAK K.  
 XX PI Han H, Kwak K;  
 XX PI WPI; 2004-707854/69.  
 XX DR N-PSDB; ADS86863.  
 XX PT Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule  
 XX PT useful for treating and/or preventing renal cachexia or inflammatory  
 XX PT cachexia.  
 XX PS Claim 13; SEQ ID NO 2; 115pp; English.  
 XX CC The present invention relates to new orthologues of human E3alpha  
 CC ubiquitin ligase, huE3alpha1 and huE3alpha1. Most intracellular proteins  
 CC are degraded through the ubiquitin-proteasome pathway. Proteins are  
 CC marked for proteasomal degradation by conjugation of ubiquitin to the  
 CC protein. Conjugation of the ubiquitin molecule involves the activation by  
 CC E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,  
 CC and then interacts with a specific E3 ligase family member. E3 ligase  
 CC binds to proteins targeted for degradation and catalyses the transfer of  
 CC ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase  
 CC determines the specificity of the system. The E3alpha family is the main  
 CC family of intracellular ligases and is involved in the N-end rule pathway  
 CC of protein degradation. E3alpha enzyme binds directly to the primary  
 CC destabilising N-terminal amino acid and catalyses ubiquitin conjugation  
 CC thereby targeting the protein for degradation. The human E3alpha gene is  
 CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-  
 CC proteasome pathway has been determined to be a major cause of rapid  
 CC muscle wasting including, fasting, metabolic acidosis, muscle  
 CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,  
 CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen  
 CC balance, burns, Cushing's syndrome, inflammatory cachexia,  
 CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-  
 CC protein malnutrition. E3alpha plays a role in the overall increase in  
 CC ubiquitination that is associated with and may mediate muscle atrophy in  
 CC catabolic and other disease states. Treatment may be administered by gene  
 CC therapy, cell therapy and antisense therapy methods. The present sequence  
 CC is human E3alpha1 protein sequence.  
 XX SQ Sequence 1749 AA;  
 Query Match 100.0%; Score 9224; DB 8; Length 1749;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MADEBAGGTERMEISAEPLQTPORLASWDDQVDFYTAFLHHLAQVLPEIYFAEMDPDLE 60

1 MADEBAGGTERMEISAEPLQTPORLASWDDQVDFYTAFLHHLAQVLPEIYFAEMDPDLE 60  
 61 KQEEVSQMSIFPTLEWYLFGEEDPDICLEKLKHSAGAFQLCGRVFKSGTETYSRCDCAIDPT 120  
 61 KQEEVSQMSIFPTLEWYLFGEEDPDICLEKLKHSAGAFQLCGRVFKSGTETYSRCDCAIDPT 120  
 121 CVLQMDCFODSVHKNHRYKQHTSTGGPCDCSDTBAWTKTGPFCVNHPEGRAGTIKENSRC 180  
 121 CVLQMDCFODSVHKNHRYKQHTSTGGPCDCSDTBAWTKTGPFCVNHPEGRAGTIKENSRC 180  
 181 PLNEBVIQVARKI PPSVIKYVVEMTIWEKEKLPELOIREKNERVYCYVFNDEHSHSYDH 240  
 181 PLNEBVIQVARKI PPSVIKYVVEMTIWEKEKLPELOIREKNERVYCYVFNDEHSHSYDH 240  
 241 VIYSQRALDCELAQAQLHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSQHPLHV 300  
 241 VIYSQRALDCELAQAQLHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSQHPLHV 300  
 301 EYLHSEIMAHQKFAIRLGSWMNKIMSYSDPQIIFCQACLRBEPDSENCLISRLMLWDA 360  
 301 EYLHSEIMAHQKFAIRLGSWMNKIMSYSDPQIIFCQACLRBEPDSENCLISRLMLWDA 360  
 361 KLYKGARKILHELI PESSPFMEYKYL PAMEFVKYKQLOKEYISDDHRSISITALSVO 420  
 361 KLYKGARKILHELI PESSPFMEYKYL PAMEFVKYKQLOKEYISDDHRSISITALSVO 420  
 421 METVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNKNFQYSDOKLGRVAVICDLK 480  
 421 METVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNKNFQYSDOKLGRVAVICDLK 480  
 481 YILISKPTIWTBLRMQFLGFRSFLKILTCMQGMEIRRQVQGHIEVDPDMEAAIAIQM 540  
 481 YILISKPTIWTBLRMQFLGFRSFLKILTCMQGMEIRRQVQGHIEVDPDMEAAIAIQM 540  
 541 QLNKNTLLMFOEWCACDEBELLVAYKECHKAVNRCTSPFISSKTVVQSCGHSLETYSRV 600  
 541 QLNKNTLLMFOEWCACDEBELLVAYKECHKAVNRCTSPFISSKTVVQSCGHSLETYSRV 600  
 601 SEDLVSIHLPLSRTLAGLHVLRLGAVSLRLEHVSFDPQVEVLVEYVLRCLVLVAQV 660  
 601 SEDLVSIHLPLSRTLAGLHVLRLGAVSLRLEHVSFDPQVEVLVEYVLRCLVLVAQV 660  
 661 AEMWRNGLSLISQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELASA 720  
 661 AEMWRNGLSLISQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELASA 720  
 721 FNKTIISTKQDQDLIKQYNTLIEBMLQVLIYIGERVYVGVNTKEEVTWREIHLICIEP 780  
 721 FNKTIISTKQDQDLIKQYNTLIEBMLQVLIYIGERVYVGVNTKEEVTWREIHLICIEP 780  
 781 MPHSAIAKNLPEVNNETGLENNVINKVATFKKPGVSGHGVYELKDESLKDFNMYFVHYSK 840  
 781 MPHSAIAKNLPEVNNETGLENNVINKVATFKKPGVSGHGVYELKDESLKDFNMYFVHYSK 840  
 841 TQHSKAEHMQKRRQKQENKDEALPPPPPEFCFAPSKVINLNCIDIMYILTVFERAID 900  
 841 TQHSKAEHMQKRRQKQENKDEALPPPPPEFCFAPSKVINLNCIDIMYILTVFERAID 900  
 901 TDSNLTWTEGMLQMAFHIALGLLEBKQLOKAPABEVTDFVHKASRLGSSAMNTQMLLE 960  
 901 TDSNLTWTEGMLQMAFHIALGLLEBKQLOKAPABEVTDFVHKASRLGSSAMNTQMLLE 960  
 961 KLKGIPOLEGQKDMITWILQMPDVTVKRLREKSLVATTSGESIKNDIETHDKKAERK 1020  
 961 KLKGIPOLEGQKDMITWILQMPDVTVKRLREKSLVATTSGESIKNDIETHDKKAERK 1020  
 1021 RKAEARLRHQKIMAQMSALQKNFIETHKLMYDNTSEMPGKEDSIWESESTPAVDYSRI 1080  
 1021 RKAEARLRHQKIMAQMSALQKNFIETHKLMYDNTSEMPGKEDSIWESESTPAVDYSRI 1080  
 1081 ALGPKRGPSVTEKEVLTCTILCOBEQVKEIENNAWVLSACVQKSTALTQHRGKPIELSGRA 1140

Db 1081 ALGPKRGPSTVEKIVLTCLIQEQEVKLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140  
QY 1141 LDPLFMDPDLAYGTTGSCGHVNHAVCWQKYFEAVQLSSQRIHVLDLFDLESGEYLCPLC 1200  
Db 1141 LDPLFMDPDLAYGTTGSCGHVNHAVCWQKYFEAVQLSSQRIHVLDLFDLESGEYLCPLC 1200  
QY 1201 KSLCNTVPIPILOPKINSEADALAOQLTLARWITQVLARI SGVNI RHAKGENPIPIF 1260  
Db 1201 KSLCNTVPIPILOPKINSEADALAOQLTLARWITQVLARI SGVNI RHAKGENPIPIF 1260  
QY 1261 FNQMGDSTLEPHSILSFVSESSIKYSNSIKEMVILFATTIYRIGLKVPDPDRPRVPM 1320  
Db 1261 FNQMGDSTLEPHSILSFVSESSIKYSNSIKEMVILFATTIYRIGLKVPDPDRPRVPM 1320  
QY 1321 TWSTCAFTQAIENLLGDEGKPLFGALQNRQNGLKALMQFAVAQRITCPQVLIOKHLVR 1380  
Db 1321 TWSTCAFTQAIENLLGDEGKPLFGALQNRQNGLKALMQFAVAQRITCPQVLIOKHLVR 1380  
QY 1381 LLSVVLPNIKSEDTPCLLSIDLPHVLVGAVALPPSLYWDDPVDLOPSSVSSSYNHLXLFH 1440  
Db 1381 LLSVVLPNIKSEDTPCLLSIDLPHVLVGAVALPPSLYWDDPVDLOPSSVSSSYNHLXLFH 1440  
QY 1441 LITMAHMLQILITVDTGLPLAQVQEDSEBAHSASSFFAEISQYTSIGSGCDIPGWLWVS 1500  
Db 1441 LITMAHMLQILITVDTGLPLAQVQEDSEBAHSASSFFAEISQYTSIGSGCDIPGWLWVS 1500  
QY 1501 LKNGITPYLRCAALPFHYLLGVTTPPELHTNAGBEYSALCSYLSLPTNLFLLFQBYWDT 1560  
Db 1501 LKNGITPYLRCAALPFHYLLGVTTPPELHTNAGBEYSALCSYLSLPTNLFLLFQBYWDT 1560  
QY 1561 VRPLLQRCADPALLNCLKQNTVVRYPKRNLSLELDPDYSCLLNQASHFRCPSADDE 1620  
Db 1561 VRPLLQRCADPALLNCLKQNTVVRYPKRNLSLELDPDYSCLLNQASHFRCPSADDE 1620  
QY 1621 RKHPVLCLFCGAILCSQNTCCOBIYNGBEVGCACIFHALHCGAGVCIFLKIRECRVVLVVG 1680  
Db 1621 RKHPVLCLFCGAILCSQNTCCOBIYNGBEVGCACIFHALHCGAGVCIFLKIRECRVVLVVG 1680  
QY 1681 KARGCAYPAPYLDEYGETDPLKRGNPLHLRSRVRKHLVWQOHCIIIEIARSQETNOM 1740  
Db 1681 KARGCAYPAPYLDEYGETDPLKRGNPLHLRSRVRKHLVWQOHCIIIEIARSQETNOM 1740  
QY 1741 LFGFNWQLL 1749  
Db 1741 LFGFNWQLL 1749

RESULT 3  
ABP58330 standard; protein; 1738 AA.

XX AC ABP58330;  
XX DT 07-APR-2003 (first entry)  
XX DE Human cell growth, differentiation and death protein CGDD-1.  
XX KW CGDD-1; cell growth; cell differentiation; cell death; human; cytostatic;  
KW antiaerobic; hepatotropic; antiinflammatory; antipsoriatic;  
KW antianemic; ophthalmological; auditory; anticonvulsant;  
KW cerebroprotective; neutropic; neuroprotective; antiparkinsonian;  
KW neuroleptic; tranquilizer; immunosuppressive; anti-HIV; antiallergic;  
KW antiasthmatic; antithyroid; antidiabetic; dermatological; nephrotropic;  
KW antirheumatic; antiarthritic; antidiabetic; antidiabetic; antidiabetic;  
KW antibacterial; fungicide; antiparasitic; protozoacide; antihelminthic;  
KW antifertility; gynaecological; ubiquitin protein ligase; enzyme;  
KW gene therapy.  
XX OS Homo sapiens.  
XX PN W0200297032-A2.  
XX XX 05-DEC-2002.

XX 05-APR-2002; 2002WO-US011152.  
XX 06-APR-2001; 2001US-0282110P.  
PR 11-APR-2001; 2001US-0283294P.  
PR 26-APR-2001; 2001US-0286820P.  
PR 27-APR-2001; 2001US-0287228P.  
PR 16-MAY-2001; 2001US-0291662P.  
PR 18-MAY-2001; 2001US-0291846P.  
PR 25-MAY-2001; 2001US-0293727P.  
PR 01-JUN-2001; 2001US-0295263P.  
PR 01-JUN-2001; 2001US-0295340P.  
PR 15-JAN-2002; 2002US-0349705P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Azimzai Y, Au-Young JK, Batra S, Baughn MR, Becha SD,  
PI Borowsky ML, Burford N, Ding L, Elliott VS, Emerling BM, Gandhi AR;  
PI Gietzen KJ, Griffin JA, Hafalia AJA, Honchell CD, Lal PG, Lee SY;  
PI Lu DAM, Arvizu CS, Ramkumar J, Reddy R, Sanjanwala MM, Tang YT;  
PI Walia NK, Wang YE, Warren BA, Xu Y, Yang J, Yao MG, Yue H;  
PI Zebajadian Y;  
XX WPI; 2003-140453/13.  
DR N-PSDB; ABZ24689.  
XX Novel human proteins associated with cell growth, differentiation and  
PT death, useful for treating, diagnosing or preventing cancer,  
PT developmental, neurological, reproductive or autoimmune/inflammatory  
PT disorders.  
XX Claim 1; Page 183-187; 238pp; English.  
XX The present sequence is the protein sequence of human CGDD-1, a novel  
CC protein associated with cell growth, differentiation and death. The  
CC sequence is predicted from Incyte clone 1351608CB1, which was isolated  
CC from a paragonomic tumour tissue cDNA library. Structural features  
CC establish the protein as being associated with cell growth,  
CC differentiation and death, with further evidence suggesting it to be a  
CC ubiquitin protein ligase. The invention is based on novel human CGDD-1 to  
CC -21 proteins (see ABP58330-50), the polynucleotides encoding them (see  
CC ABZ24689-709), and to the use of these for the diagnosis, treatment or  
CC prevention of cell proliferative disorders including cancer,  
CC developmental disorders, neurological disorders, autoimmune disorders,  
CC reproductive disorders, and disorders of the placenta, and in the  
CC assessment of the effects of exogenous compounds on the activity and  
CC expression of proteins and nucleic acids associated with cell growth,  
CC differentiation and death  
XX SQ Sequence 1738 AA;  
Query Match 99.4%; Score 9168; DB 6; Length 1738;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 MEISAEPLQTPORLASWDDQVDFYAFHLHLAQLVPELYPAEMDPDLKQESVQMSIF 71  
Db 1 MEISAEPLQTPORLASWDDQVDFYAFHLHLAQLVPELYPAEMDPDLKQESVQMSIF 60  
QY 72 TPLEWYLFGEHPDILCLEKLKHSAGFQLCGRVPKSGGETTYSKRDCAIDPTCVLCMDCFQDS 131  
Db 61 TPLEWYLFGEHPDILCLEKLKHSAGFQLCGRVPKSGGETTYSKRDCAIDPTCVLCMDCFQDS 120  
QY 132 VHKHRYKMHSTTGGGFCGCGTEAWKTGPPFCVNHPEGRAGTIKNSRCPLNEEIVQAR 191  
Db 121 VHKHRYKMHSTTGGGFCGCGTEAWKTGPPFCVNHPEGRAGTIKNSRCPLNEEIVQAR 180  
QY 192 KIFPSPVIVYVEMTWEEKEKLPPELQIREKNERYCYLVFNDEHHSYDHVYLSQALDC 251  
Db 181 KIFPSPVIVYVEMTWEEKEKLPPELQIREKNERYCYLVFNDEHHSYDHVYLSQALDC 240  
QY 252 ELAEALQHTTATDKGRRRAVKAAGAACQAEKEDIKSHSENVYSQPLHVEVLHSEIMAHQ 311

Db 241 ELAEALHTTAIDKEGRAVAGAYAACQAEKEDIKSHSENVSOHPLHVEVLHSEIMAHQ 300  
Qy 312 KPALRLGSMWNKIMSYSSDFRQIFCOACIARBPDSNPCLISRLMLWDALYKGAARKILH 371  
Db 301 KPALRLGSMWNKIMSYSSDFRQIFCOACIARBPDSNPCLISRLMLWDALYKGAARKILH 360  
Qy 372 ELIFSSFFMEMEYKULPAMEFVKYKOLQKEYISDDHDSISITALSVMQFTVPTLARHL 431  
Db 361 ELIFSSFFMEMEYKULPAMEFVKYKOLQKEYISDDHDSISITALSVMQFTVPTLARHL 420  
Qy 432 IBEQNVISVITETLLEVLPEYLDNRNKNFQGSQDKLGRVAVICDLKYILISKPTIWT 491  
Db 421 IBEQNVISVITETLLEVLPEYLDNRNKNFQGSQDKLGRVAVICDLKYILISKPTIWT 480  
Qy 492 ERLRMQFLGFRSFLKILTCMQGMBEIRQVQGHIEVDPDWEAAIAIQMOLKNILLMFOE 551  
Db 481 ERLRMQFLGFRSFLKILTCMQGMBEIRQVQGHIEVDPDWEAAIAIQMOLKNILLMFOE 540  
Qy 552 WCACDELLLVAYKECHKAVMRCSFISSTKTVOGSHLETYSYRVEDLSVTHLPL 611  
Db 541 WCACDELLLVAYKECHKAVMRCSFISSTKTVOGSHLETYSYRVEDLSVTHLPL 600  
Qy 612 SRTLAGLHVLRLGAVSRILHBEVSPEDQVEVLVEYPLRCLVLAQVVAEMWRRNGLSL 671  
Db 601 SRTLAGLHVLRLGAVSRILHBEVSPEDQVEVLVEYPLRCLVLAQVVAEMWRRNGLSL 660  
Qy 672 ISQVFYQDVKREEMYDKDIIIMLOIGASLMDPNKPLLLVLQRYELAEAFNKTIISTKQD 731  
Db 661 ISQVFYQDVKREEMYDKDIIIMLOIGASLMDPNKPLLLVLQRYELAEAFNKTIISTKQD 720  
Qy 732 LIKQVNTLIEEMLOVLIYVGERYVPGVGNVTKKEEVTMBEIIHLLCIEPMPSIAKNLP 791  
Db 721 LIKQVNTLIEEMLOVLIYVGERYVPGVGNVTKKEEVTMBEIIHLLCIEPMPSIAKNLP 780  
Qy 792 ENENNETGLENVINKVATPKKGVSGHGVYELKDBSLKDPNMYFHYSTQHSKARHMQ 851  
Db 781 ENENNETGLENVINKVATPKKGVSGHGVYELKDBSLKDPNMYFHYSTQHSKARHMQ 840  
Qy 852 KRRKQENKDEALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLTGML 911  
Db 841 KRRKQENKDEALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLTGML 900  
Qy 912 QNAFHILALGLLEEKQOLQKABEEVTFPYHKASRLGSSANNIOMLKLKGIPOLEGQ 971  
Db 901 QNAFHILALGLLEEKQOLQKABEEVTFPYHKASRLGSSANNIOMLKLKGIPOLEGQ 960  
Qy 972 KDMITWILQMEDTVKRLKSKCLIVATTSSGSEIKNDEITHDKKAEARKKAEARLHRQ 1031  
Db 961 KDMITWILQMEDTVKRLKSKCLIVATTSSGSEIKNDEITHDKKAEARKKAEARLHRQ 1020  
Qy 1032 KIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEEESTPAVSDYSRIALGPGRGPSVT 1091  
Db 1021 KIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEEESTPAVSDYSRIALGPGRGPSVT 1080  
Qy 1092 EKEVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDIA 1151  
Db 1081 EKEVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDIA 1140  
Qy 1152 YGTYTGSCGHVHVCWQKYFEAVOLSSOORLHVLDLFDLESGEYLCPLCKSCNTVPIPI 1211  
Db 1141 YGTYTGSCGHVHVCWQKYFEAVOLSSOORLHVLDLFDLESGEYLCPLCKSCNTVPIPI 1200  
Qy 1212 PLQPKINSENADALAQLLTLARWIQTVLARISGVNIRHAKGENPIPIFNQMGDSTLE 1271  
Db 1201 PLQPKINSENADALAQLLTLARWIQTVLARISGVNIRHAKGENPIPIFNQMGDSTLE 1260  
Qy 1272 PHSILSPGVESIKYSNSIKEMVILPATTIYRIGLKVPPDERDPRVPMLTWTSCAPTIOA 1331  
Db 1261 PHSILSPGVESIKYSNSIKEMVILPATTIYRIGLKVPPDERDPRVPMLTWTSCAPTIOA 1320  
Qy 1332 IENILGDEGKPLFGALONRQHGLKALMQPAVAORITCQVLIQKHLVRLLSVLPNKS 1391  
Db 1321 IENILGDEGKPLFGALONRQHGLKALMQPAVAORITCQVLIQKHLVRLLSVLPNKS 1380

## RESULT 4

ADJ95459

ID ADJ95459 standard; protein; 1734 AA.

XX AC ADJ95459;

XX DT 03-JUN-2004 (first entry)

XX Human Ubiquitin ligase E3alpha I variant.

XX Human; enzyme; ubiquitin ligase; E3alpha I; ubiquitin-proteasome pathway;  
gene therapy; vaccine; muscular atrophy; cachexia; catabolic disorders;  
cancer cachexia; renal cachexia; inflammatory cachexia;  
muscle wasting disorder; metabolic acidosis; uremia; burn;  
hyperthyroidism; Cushing's syndrome; fasting; denervation atrophy;  
diabetes mellitus; sepsis; AIDS wasting syndrome; SNP;  
single nucleotide polymorphism.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 225..246 /note= "Encoded by nucleotides 658-678 of SEQ ID 18 "

FT Misc-difference 1573 /note= "May be Tip as the result of a single nucleotide

FT polymorphism"

FT Misc-difference 1665..1666 /note= "Encoded by nucleotides 4933-4998 of SEQ ID 18 "

XX US6706505-B1.

XX 16-MAR-2004.

XX 28-NOV-2000; 2000US-00724126.

XX 08-MAR-2000; 2000US-0187911P.

XX (AMGE-) AMGEN INC.

XX Han H, Kwak K;

XX WPI; 2004-236723/22.

XX N-PSDB; ADJ95458.

XX New nucleic acid molecule, useful for preparing a composition for  
diagnosing, treating or preventing diseases associated with human





Db	1726	LFGENWQLL 1734	
RESULT 5			
ADSR6881			
ID	ADSR6881	standard; protein; 1734 AA.	
XX	AC	ADSR6881;	
XX	DT	16-DEC-2004 (first entry)	
XX	DE	Variant human E3alpha1 ubiquitin ligase.	
XX	KW	Human; E3alpha ubiquitin ligase; huE3alpha1;	
XX	KW	ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;	
XX	KW	fasting; metabolic acidosis; muscle degeneration; kidney failure;	
XX	KW	renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;	
XX	KW	cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;	
XX	KW	inflammatory cachexia; hyperthyroidism; denervation atrophy;	
XX	KW	protein/tissue wasting; energy-protein malnutrition; muscle atrophy;	
XX	KW	gene therapy; enzyme.	
XX	OS	Homo sapiens.	
XX	PN	US2004185037-A1.	
XX	PD	23-SEP-2004.	
XX	PF	15-JAN-2004; 2004US-00758672.	
XX	PR	08-MAR-2000; 2000US-0187911P.	
XX	PR	28-NOV-2000; 2000US-00724126.	
XX	PA	(HANH/) HAN H.	
XX	PA	(KWAK/) KWAK K.	
XX	PI	Han H, Kwak K;	
XX	DR	WPI; 2004-707854/69.	
XX	PT	Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule	
XX	PT	useful for treating and/or preventing renal cachexia or inflammatory	
XX	PT	cachexia.	
XX	PS	Example 9; SEQ ID NO 19; 115pp; English.	
XX	CC	The present invention relates to new orthologues of human E3alpha	
XX	CC	ubiquitin ligase, huE3alpha1 and huE3alpha1. Most intracellular proteins	
XX	CC	are degraded through the ubiquitin-proteasome pathway. Proteins are	
XX	CC	marked for proteasomal degradation by conjugation of ubiquitin to the	
XX	CC	protein. Conjugation of the ubiquitin molecule involves the activation by	
XX	CC	E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,	
XX	CC	and then interacts with a specific E3 ligase family member. E3 ligase	
XX	CC	binds to proteins targeted for degradation and catalyses the transfer of	
XX	CC	ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase	
XX	CC	determines the specificity of the system. The E3alpha family is the main	
XX	CC	family of intracellular ligases and is involved in the N-end rule pathway	
XX	CC	of protein degradation. E3alpha enzyme binds directly to the primary	
XX	CC	destabilising N-terminal amino acid and catalyses ubiquitin conjugation	
XX	CC	thereby targeting the protein for degradation. The human E3alpha gene is	
XX	CC	located on chromosome 15 q. Increased proteolysis through the ubiquitin-	
XX	CC	proteasome pathway has been determined to be a major cause of rapid	
XX	CC	muscle wasting including, fasting, metabolic acidosis, muscle	
XX	CC	degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,	
XX	CC	sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen	
XX	CC	balance, burns, Cushing's syndrome, inflammatory cachexia,	
XX	CC	hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-	
XX	CC	protein malnutrition. E3alpha plays a role in the overall increase in	
XX	CC	ubiquitination that is associated with and may mediate muscle atrophy in	
XX	CC	catabolic and other disease states. Treatment may be administered by gene	
XX	CC	therapy, cell therapy and antisense therapy methods. The present sequence	
XX	CC	is a variant of human E3alpha1 ubiquitin ligase. The patent describes	
XX	CC	this sequence as containing a Trp at position 1568 as a result of a SNP	
CC	in the nucleotide sequence. However, the sequence contains a Pro at		
CC	position 1568.		
XX	CC	Sequence 1734 AA;	
SQ	Query Match	98.5%; Score 9087; DB 8; Length 1734;	
	Best Local Similarity	98.8%; Pred. No. 0;	
	Matches 1728; Conservative	0; Mismatches 1; Indels 20; Gaps 1;	
QY	1	MADBEAGGTGTERMEISAELEPQTQRLASWMDQVDFYAFLLHLAQLVPEIYFAEMDPDLE	60
DB	6	MADBEAGGTGTERMEISAELEPQTQRLASWMDQVDFYAFLLHLAQLVPEIYFAEMDPDLE	65
QY	61	KQESVQMSIFPLEWYLFGEDEPDI CLEKLKUSGAFOLCGRVFKSGGETTYSRCDCAIDPT	120
DB	66	KQESVQMSIFPLEWYLFGEDEPDI CLEKLKUSGAFOLCGRVFKSGGETTYSRCDCAIDPT	125
QY	121	CVLGMDCTQDSVHKVHRYKMTSTGGFCDCGDTBAWTKGPPCVNHEPGRAGTICKENSRC	180
DB	126	CVLGMDCTQDSVHKVHRYKMTSTGGFCDCGDTBAWTKGPPCVNHEPGRAGTICKENSRC	185
QY	181	PLNEEVIIVQARKIPPSVIKYVVMETIWBEEKELPPELQIREKNERYCYVLFNDEHHSYDH	240
DB	186	PLNEEVIIVQARKIPPSVIKYVVMETIWBEEKELPPELQIREKNERYCYVLFNDEHHSYDH	245
QY	241	VIYSLQALDCELAQAQLHTTAIDKEGRRVAVAGAYAAQCAKEDI KHSERNVSORPLHV	300
DB	246	VIYSLQALDCELAQAQLHTTAIDKEGRRVAVAGAYAAQCAKEDI KHSERNVSORPLHV	305
QY	301	EVLHSEINAHOKPALRLGSMWNKINSYSDPRQIFCOACLRREPDPENPCLLSRMLWDA	360
DB	306	EVLHSEINAHOKPALRLGSMWNKINSYSDPRQIFCOACLRREPDPENPCLLSRMLWDA	365
QY	361	KLYGKARKILHELIFSSPFMEYKCLFAMEPVKYYKQLQKEYISDDHDSISITALSQ	420
DB	366	KLYGKARKILHELIFSSPFMEYKCLFAMEPVKYYKQLQKEYISDDHDSISITALSQ	425
QY	421	MFTVPTLARHLIEBONVISVITETLLVLPEVLDNNKPNFGYSGDKLGRVAVICDLK	480
DB	426	MFTVPTLARHLIEBONVISVITETLLVLPEVLDNNKPNFGYSGDKLGRVAVICDLK	485
QY	481	YILSKPTIWTERRMQLEGFRSEKILTCMQGMEIRROVQGHIVDPDHEAIAIOM	540
DB	486	YILSKPTIWTERRMQLEGFRSEKILTCMQGMEIRROVQGHIVDPDHEAIAIOM	545
QY	541	QLKNILLMFQWCACDEBLLVAYKECHAVNRCSTSPISSTKVQSCGHSLETYSYRV	600
DB	546	QLKNILLMFQWCACDEBLLVAYKECHAVNRCSTSPISSTKVQSCGHSLETYSYRV	605
QY	601	SEDLVSIHLPLSRTLAGLHVLRLSRIGAVSRLEHFPVSFEDQVEVLVYPLRCLVLVAQV	660
DB	606	SEDLVSIHLPLSRTLAGLHVLRLSRIGAVSRLEHFPVSFEDQVEVLVYPLRCLVLVAQV	665
QY	661	AMWRRNGLSISQVFFYQDYKCREMYDKDIIMLQIGASLMDPNKFLLLVQLRYELAE	720
DB	666	AMWRRNGLSISQVFFYQDYKCREMYDKDIIMLQIGASLMDPNKFLLLVQLRYELAE	725
QY	721	FNKTIKQDLIKQYNTLIEBMLQVLIYVGERVPGVGNVYKEVTRREIHLILCTEP	780
DB	726	FNKTIKQDLIKQYNTLIEBMLQVLIYVGERVPGVGNVYKEVTRREIHLILCTEP	785
QY	781	MPHSAIAKNLPENENNETGLENVINKVATFKKPGVSGHVGVELKDESKDFNMFYHYSK	840
DB	786	MPHSAIAKNLPENENNETGLENVINKVATFKKPGVSGHVGVELKDESKDFNMFYHYSK	845
QY	841	TOHSAEIMQKRRKQENKDBALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAID	900
DB	846	TOHSAEIMQKRRKQENKDBALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAID	905
QY	901	TDSNLTWTEGMLQMAFHIALGILLBEKQOLQKAPBEVTFDFYHKASRLGSSAMNQMLE	960
DB	906	TDSNLTWTEGMLQMAFHIALGILLBEKQOLQKAPBEVTFDFYHKASRLGSSAMNQMLE	965

QY 961 KLGIPQLEGQDMITWILQMFDTVKRLREKSLVATVTSSESINKDEITHDKEARK 1020  
DB |||||  
QY 966 KLGIPQLEGQDMITWILQMFDTVKRLREKSLVATVTSSESINKDEITHDKEARK 1025  
DB |||||  
QY 1021 RKAARLHRQKIMQMSALQNFIFETHKLMYDNTSEMPGKEDSIWEEESTPAVDSYRI 1080  
DB |||||  
QY 1026 RKAARLHRQKIMQMSALQNFIFETHKLMYDNTSEMPGKEDSIWEEESTPAVDSYRI 1085  
DB |||||  
QY 1081 ALGPKRGPSVTEKVLTCILQEBQEVKIENNAMVLSACVQKSTALTQHRGHPIELSGBA 1140  
DB |||||  
QY 1086 ALGPKRGPSVTEKVLTCILQEBQEVKIENNAMVLSACVQKSTALTQHRGHPIELSGBA 1145  
DB |||||  
QY 1141 LDPLFMDPLAYGTGTGSGHVMHACWQKYPEAVQLSSQQRIHVLDLPLESGEYLCPIC 1200  
DB |||||  
QY 1201 KSLCNTVPIIPILOPKINSENADALAOILLTARWITQTVLARIISGYNIRHAKGENPIPIF 1260  
DB |||||  
QY 1261 FNQMGDSITLPHSILSPGVSSIKYSIKEMVILFATTIYRIGLVKVPDPDRPRVPM 1320  
DB |||||  
QY 1266 FNQMGDSITLPHSILSPGVSSIKYSIKEMVILFATTIYRIGLVKVPDPDRPRVPM 1325  
DB |||||  
QY 1321 TWSTCAFTIQAENILGDEKPLFGALQNRHNGKALMOFAVQRIITCPQVLIQKHLVR 1380  
DB |||||  
QY 1326 TWSTCAFTIQAENILGDEKPLFGALQNRHNGKALMOFAVQRIITCPQVLIQKHLVR 1385  
DB |||||  
QY 1381 LLSVLPNPKSBDTPCLLSIDLPHVLGAVLAPPSLYWDDPVDLPSSVSSSYNHLFLH 1440  
DB |||||  
QY 1441 LITMAHMLQILLTVDLPLAQVQEDSEAHASFFABISQYTSIGCDIPGWYLVMS 1500  
DB |||||  
QY 1501 LKNGITPYLRCAALPHYLGLVTPPEELHTNSAEGEYSALCSYLSIPTNLFLLPQBYMDT 1560  
DB |||||  
QY 1561 VRPLLQRCADPALLNCLKQNTVVRYPKRNLSIELPDDYSCLNQASHFRCPSADDE 1620  
DB |||||  
QY 1621 RHPVLCPLCGAILCSQNICQEI VNGEVGACIFHALHCGAGVCI FZKIRECRVVLVEG 1680  
DB |||||  
QY 1681 KARGCAYPAPYLDYGETDPLKRGNPPLHLSBERYKHLVWQHCIIIEIARSQETNOM 1740  
DB |||||  
QY 1741 LFGFNWOLL 1749  
DB |||||  
QY 1726 LFGFNWOLL 1734  
DB |||||

RESULT 6  
ID AAW84351  
XX AAW84351 standard; protein; 1757 AA.  
AC AAW84351;  
XX  
DT 25-MAR-1999 (first entry)  
DE Murine ubiquitin-protein ligase Ub1.  
XX Ubiquitin-protein ligase; Ub1; mouse; ubiquitinylation; degradation;  
KW N-end rule pathway; stress-related muscle wasting; inhibitor; screen.  
XX Mus sp.  
XX  
PN US5861312-A.

XX 19-JAN-1999.  
PD  
XX 02-DEC-1997; 97US-00982956.  
PF  
XX 02-DEC-1997; 97US-00982956.  
PR  
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
PA  
XX Kwon YT, Varshavsky A;  
PI WPI; 1999-130395/11.  
DR N-PSDB; AAV99308.  
XX Mouse and human Ub1 cDNA - useful for producing recombinant Ub1  
PT polypeptides.  
XX Disclosure; Col 15-28; 18pp; English.  
XX The present sequence represents a ubiquitin-protein ligase called Ub1.  
CC The Ub1 enzymes are involved in protein ubiquitinylation and ultimate  
CC degradation through the N-end rule pathway and have been linked to stress  
CC -related muscle wasting. Recombinant Ub1 polypeptides can be used to  
CC screen for inhibitors of muscle wasting when this is associated with the  
CC N-end rule pathway  
XX Sequence 1757 AA;  
SQ  
Query Match 93.4%; Score 8617; DB 2; Length 1757;  
Best Local Similarity 92.1%; Pred. NO. 0;  
Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;  
QY 1 MADEAGTHERMEISAEIPQTPORLASWMDQVDFYTAFLHLAQLVPEIYFAEMDPLDLE 60  
DB |||||  
QY 61 KQESVQMSIITPLEWYLFGBDDPDIKLEKXHSAGFOLCGRVKFSGETTYSRCDCAIDPT 120  
DB |||||  
QY 61 KQESVQMSIITPLEWYLFGBDDPDIKLEKXHSAGFOLCGRVKFSGETTYSRCDCAIDPT 120  
DB |||||  
QY 121 CVLCMDCFQDSVHKHRYKMTSTGGFCDCDGTAEWTKGFCVNVHPEGRAGTIKENSRC 180  
DB |||||  
QY 121 CVLCMDCFQDSVHKHRYKMTSTGGFCDCDGTAEWTKGFCVNVHPEGRAGTIKENSRC 180  
DB |||||  
QY 181 PLNEEVIQARRIIPPSVIKYIVEMTIWEEKELPELOIREKNERYCYVLFNDEHSHYDH 240  
DB |||||  
QY 181 PLNEEVIQARRIIPPSVIKYIVEMTIWEEKELPELOIREKNERYCYVLFNDEHSHYDH 240  
DB |||||  
QY 241 VYLSQALDCELAQAOLHTTAIDKEGRVAKAGAYAAACQAKEDIKSHSENVSOHPLHV 300  
DB |||||  
QY 241 VYLSQALDCELAQAOLHTTAIDKEGRVAKAGAYAAACQAKEDIKSHSENVSOHPLHV 300  
DB |||||  
QY 301 EYLHSEIMAHQKFPALRLGSMWNKIMSYSDFQIIFCQACLREPDSENPCLISRLMLWDA 360  
DB |||||  
QY 301 EYLHSEIMAHQKFPALRLGSMWNKIMSYSDFQIIFCQACLREPDSENPCLISRLMLWDA 360  
DB |||||  
QY 361 KLYGARKILHELIFSPFMEMEYKGLFAMEVVKYKQLQKEYISDDHRSISITALSVO 420  
DB |||||  
QY 361 KLYGARKILHELIFSPFMEMEYKGLFAMEVVKYKQLQKEYISDDHRSISITALSVO 420  
DB |||||  
QY 421 MPTVPTLARHLIEEONVISVITETLLEVLPEYLDNRNKNFNFQYSGQDKLGRVYAVICDLK 480  
DB |||||  
QY 421 MPTVPTLARHLIEEONVISVITETLLEVLPEYLDNRNKNFNFQYSGQDKLGRVYAVICDLK 480  
DB |||||  
QY 481 YLISKPTIWTTERLSMQFLEGFRSPLKILTCMQGMEIRROVQGHIEVDPDWEAAIAIQM 540  
DB |||||  
QY 481 YLISKPTIWTTERLSMQFLEGFRSPLKILTCMQGMEIRROVQGHIEVDPDWEAAIAIQM 540  
DB |||||  
QY 541 QLNKILLMFQEWCADEDELLLVAYKECHKAVNRCSTSTISSKTVVQSGHSLKTSYKV 600  
DB |||||  
QY 541 QLNKILLMFQEWCADEDELLLVAYKECHKAVNRCSTSTISSKTVVQSGHSLKTSYKV 600  
DB |||||  
QY 601 SEDLVSIHPLSRTLAGLHVRLSGAVSRLEHFVSFDPQVEVLVEYVLEPCLVLVAQVY 660  
DB |||||

601 SEDVSIHLPLSRTLAGLHVRSLGAIISRLHEFPDFDSFQVEVLVEYFLRCLVLVAQVV 660  
661 AEMWRNGLSLISQVYQDVKREEMDKDIIIMQIGASLMDPNKFLLLVLQRYELAEA 720  
661 AEMWRNGLSLISQVYQDVKREEMDKDIIIMQIGASIMDPNKFLLVLQRYELTDA 720  
721 FNKTIKTDODLIIKQNTLIEBMLQVLIYIVGERYVPGVGNVTKBERVWREIILHLCIEP 780  
721 FNKTIKTDODLIIKQNTLIEBMLQVLIYIVGERYVPGVGNVTKBERVWREIILHLCIEP 780  
781 MPHSAIAKLIPENNETGLENVINKVATPKPGVSGHGVYELKDSLADFNMYFYHSK 840  
781 MPHSAIARLIPENNETGLENVINKVATPKPGVSGHGVYELKDSLADFNMYFYHSK 840  
841 TQSKAEHMKRRKQENKDEALPPPPPEFCFAFSKVINLNCIDIMMYILRTVFERAID 900  
841 TQSKAEHMKRRKQENKDEALPPPPPEFCFAFSKVINLNCIDIMMYILRTVFERAID 900  
901 TDSNLWTEGMLQWAFHILALGLLEBKQLOKAPBEVTPDFTHKASRLGSSAM---NIQM 957  
901 TESNLWTEGMLQWAFHILALGLLEBKQLOKAPBEVTPDFTHKASRLGSSAMNAQIOM 960  
958 LLEKLGIPQSGQDMITWILQMDTVKRLREKSLIVATTSGSESIKNDSEITHDKKA 1017  
961 LLERLKGIPQSGQDMITWILQMDTVKRLREKSLIVATTSGLECIKSEITHDKKA 1020  
1018 ERKKAERARLHROKIMQMSALQKNFIETHKMTDNTSEMPCKEDSIMEESTPAVS DY 1077  
1021 ERKKAERARLHROKIMQMSALQKNFIETHKMTDNTSEVTKEDSIMEESTPAVSEA 1080  
1078 SERIALGPKRGPVTEKEVLTCLIQEORVKIENNAMVLSACVQKSTALTQHRGPIELS 1137  
1081 SERIALGPKRGPVTEKEVLTCLIQEORVKIENNAMVLSACVQKSTALTQHRGKPDHL 1140  
1138 GEALDPLENDPLAGVTYTGSGHVMHVCWKYFEANVQLSSQORIHVDLPLEGEVILC 1197  
1141 GETLOPLFMDPLAGVTYTGSGHVMHVCWKYFEANVQLSSQORIHVDLPLEGEVILC 1200  
1198 PLCKSLCNTVPIIPILOPKINSENADALAQLLTLARWITQTLARISGYNIRHAKGENP- 1256  
1201 PLCKSLCNTVPIIPILOPKINSENADALAQLLTLARWITQTLARISGYNIRHAKGEAPA 1260  
1257 IPFPNQMGDSTLFHSLSGVSSIKSYNSIKEMWILFATTIYRIGLKVPDPDR 1316  
1261 VPVLFNQMGDSTLFHSLSGVSSIKSYNSIKEMWILFATTIYRIGLKVPDPDELDR 1320  
1317 VPMLTWSTCAFTIOAIENLLGDEKPLFGALQNRHNGIKALMQFAVAQRITCPQVLIQK 1376  
1321 VPMTWTSTCAFTIOAIENLLGDEKPLFGALQNRHNGIKALMQFAVAQRATCPQVLIHK 1380  
1377 HLVRLLSVLVPNIKSEDTPCLLSIDLPHVLGVALAPSLYWDVDPDLPSSVSSYNHL 1436  
1381 HUALRLSVLPNKGENTPGLLSVDLPHVLGVALAPSLYWDVDPDLPSSVSSYNHL 1440  
1437 YLFHLITWMLQILLTVDVT---GLPLAQVEDSEEAHSFPFAEISQYTSIGCDI 1492  
1441 YLFHLITWMLQILLTVDLSPGPPLAEGEEDSEARCAAPFVEVSQHTDGLTGCA 1500  
1493 PGWYLVWSLXNGITPYLRCAALFPHYLGVTTPPELHTNSARGEVSALCSYLSLTPNLFL 1552  
1501 PGWYLVWSLXNGITPYLRCAALFPHYLGVTTPPELHTNSARGEVSALCSYLSLTPNLFL 1560  
1553 LPQEWYDVTVRPLQWRCADPALLNCLIKONTVVRYPKENSILIEPDDYSCLLNQAHSFR 1612  
1561 LPQEWYDVTIRPLQWRCADPALLNCLIKONTVVRYPKENSILIEPDDYSCLLNQAHSFR 1620  
1613 CPRSADDERKHPVLCIFCGAILCSQNICQOEIVNGEEVGACIFHALHCGAGVCIPLKIRE 1672  
1621 CPRSADDERKHPVLCIFCGAILCSQNICQOEIVNGEEVGACVFFHALHCGAGVCIPLKIRE 1680  
1673 CRVVLVEGKARGCAYPAPYLDYGETDPGLKRGNPLHLSRERYRKLHLVWQOHCIIIBETA 1732

Db 1681 CRVVLVEGKARGCAYPAPYLDYGETDPGLKRGNPLHLSRERYRKLHLVWQOHCIIIBETA 1740  
QY 1733 RSQETNQLMFGFNWOLL 1749  
Db 1741 RSQETNQLMFGFNWOLL 1757  
RESULT 7  
ID AAB31162 standard, protein; 1757 AA.  
XX AC AAB31162;  
XX DT 02-APR-2001 (first entry)  
XX Amino acid sequence of Mouse Ubr1 protein.  
XX Ubr1; E3-type protein; ubiquitin system; ubiquitin-protein ligase;  
XX N-end rule pathway; intracellular pathogen; Lysteria monocytogenes;  
XX Yersinia enterocolitica; muscle wasting; infection.  
XX Mus sp.  
XX US6159732-A.  
XX PD 12-DEC-2000.  
XX PF 11-JAN-1999; 99US-00228317.  
XX PR 02-DEC-1997; 97US-00982956.  
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
XX Kwon YT, Varshavsky A;  
DR WPI. 2001-090278/10.  
DR N-PSDB; AAC86933.  
XX Inhibiting the N-end rule pathway in mammalian cells for treating  
infections and various diseases associated with muscle tissue wasting, by  
inhibiting the expression of Ubr1 gene.  
XX Example; Col 15-28; 18pp; English.  
XX The present sequence represents a murine Ubr1 enzyme. Ubr1 is an E3-type  
protein of the ubiquitin system. Specifically, it is a ubiquitin-protein  
ligase. The enzyme is specific for destabilising residues exposed at the  
N-terminus of protein substrates. Inhibition of the expression of Ubr1  
gene in a cell results in inhibition of the N-end rule pathway. The  
method is used for treatment of mammalian cells infected with an  
intracellular pathogen, e.g. Lysteria monocytogenes or Yersinia  
enterocolitica. Inhibition of N-end rule pathway is also useful for  
treating various diseases associated with wasting of muscle tissue and  
infections  
XX SQ Sequence 1757 AA;  
Query Match 93.4%; Score 8617; DB 4; Length 1757;  
Best Local Similarity 92.1%; Pred. No. 0;  
Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;  
QY 1 MADBEAGTERMEISAEPLPOTFQRLASWDDQVDFYTAFLHHLAQLVPEIYFAEMDPDLE 60  
Db 1 MADBEAGTERMEISAEPLPOTFQRLASWDDQVDFYTAFLHHLAQLVPEIYFAEMDPDLE 60  
QY 61 KOESVQMSITPLEWYLFGEEDPDICLEKXHSAPQLCGRVFKSGGTTTSCRDCAIDPT 120  
Db 61 KOESVQMSITPLEWYLFGEEDPDICLEKXHSAPQLCGRVFKSGGTTTSCRDCAIDPT 120  
QY 121 CVLCHDCQDVSVHNHRYKMTSTGGGFCDCGDEATKTPPCVNHPEGRAGTIKENSRC 180  
Db 121 CVLCHDCQDVSVHNHRYKMTSTGGGFCDCGDEATKTPPCVNHPEGRAGTIKENSRC 180  
121 CVLCHDCQSSVHNHRYKMTSTGGGFCDCGDEATKTPPCVNHPEGRAGTIKENSRC 180

QY 181 PLNEEVIOARKIPFSVIVVEMTIBEEKELPELQIREKNERYCYVLFNDEHSHYD 240  
DB 181 PLNEEVIAQARRIFPSVIKYIEMTIBEEKELPELQIREKNERYCYVLFNDEHSHYD 240  
QY 241 VIYSIQRALDCLBAEQALHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSHPLHV 300  
DB 241 VIYSIQRALDCLBAEQALHTTAIDKEGRRAVAGAYATCQEAKEEDIKSHSENVSHPLHV 300  
QY 301 EYLHSEIWAHQFALRIASWNNKINSYSDFRQIFCOACLEREPPSENPCLISRLMLWDA 360  
DB 301 EYLHSEIWAHQFALRIASWNNKINSYSDFRQIFCOACLEREPPSENPCLISRLMLWDA 360  
QY 361 KLYKGARKILHBLIFSSPFMEYKCLFAMEFVKYKLOKEYISDDHRSISITALSVO 420  
DB 361 KLYKGARKILHBLIFSSPFMEYKCLFAMEFVKYKLOKEYISDDHRSISITALSVO 420  
QY 421 MFTVPTLARHLIEQNVISVITETILLEVLPEYLDNRNKNFQSGYQDKLGRYAVICDLK 480  
DB 421 MFTVPTLARHLIEQNVISVITETILLEVLPEYLDNRNKNFQSGYQDKLGRYAVICDLK 480  
QY 481 YLISKPTITWELRMQFLEGFRSFLKILTCQGHBEIRROVGOHIEVDPDWEAAIAIOM 540  
DB 481 YLISKPTITWELRMQFLEGFRSFLKILTCQGHBEIRROVGOHIEVDPDWEAAIAIOM 540  
QY 541 QLNKILLMPOEWCACDEBLLIVAYKECHKAVNRCTSPISSTKVVSQSGHSLKTSYRV 600  
DB 541 QLNKILLMPOEWCACDEBLLIVAYKECHKAVNRCTSPISSTKVVSQSGHSLKTSYRV 600  
QY 601 SEDLVSIIHLPLSRTLAGHLVRLSRLGAVSRLEHFEVSPEDQVEVLPEYLRCLVLVAQV 660  
DB 601 SEDLVSIIHLPLSRTLAGHLVRLSRLGAVSRLEHFEVSPEDQVEVLPEYLRCLVLVAQV 660  
QY 661 AEMWRNGLSLISQVFTYQDVKCRBEMDKOIIIMLOIGASLMDPNKPLLLVLQRYELABA 720  
DB 661 AEMWRNGLSLISQVFTYQDVKCRBEMDKOIIIMLOIGASLMDPNKPLLLVLQRYELABA 720  
QY 721 FNKTIISTKDQDILIKOYNTLIEBMLQVLIYIGERVVPGVGNVTKBEVTWREIITHLCIEP 780  
DB 721 FNKTIISTKDQDILIKOYNTLIEBMLQVLIYIGERVVPGVGNVTKBEVTWREIITHLCIEP 780  
QY 781 MPHSAIAKNLPENNETGLENNVINKVATFKKPGVSGHGVYELKDESLDKFNMYFYHSK 840  
DB 781 MPHSAIAKNLPENNETGLENNVINKVATFKKPGVSGHGVYELKDESLDKFNMYFYHSK 840  
QY 841 TQHSKAEHMQKRRQKQENKDEALPPPPPPPCFAPSKVINLNCIMIMYILRTVPERAID 900  
DB 841 TQHSKAEHMQKRRQKQENKDEALPPPPPPPCFAPSKVINLNCIMIMYILRTVPERAID 900  
QY 901 TDSNLTWTEGMLQWAFHILALGLEKQLOKQAPBEVTPDFVHKASRLGSSAM--NIQM 957  
DB 901 TDSNLTWTEGMLQWAFHILALGLEKQLOKQAPBEVTPDFVHKASRLGSSAM--NIQM 957  
QY 958 LLEKLGKIPQLEGQKDMITWILQMFDTVRLREKSLIVATTSGSEIKNDEIITHDKKA 1017  
DB 958 LLEKLGKIPQLEGQKDMITWILQMFDTVRLREKSLIVATTSGSEIKNDEIITHDKKA 1017  
QY 1018 ERKRAEAARLHRQKIMQMSALQKNPIETHKMLDNTSEMPGKEDSIMEBESTPAVS DY 1077  
DB 1018 ERKRAEAARLHRQKIMQMSALQKNPIETHKMLDNTSEMPGKEDSIMEBESTPAVS DY 1077  
QY 1078 SERIALGPKGSPVTEKEVLTCILCOBEQVKIENNAMVLSACVOKSTALTQHRGKPIELS 1137  
DB 1078 SERIALGPKGSPVTEKEVLTCILCOBEQVKIENNAMVLSACVOKSTALTQHRGKPIELS 1137  
QY 1138 SERIALGPKGSPVTEKEVLTCILCOBEQVKIENNAMVLSACVOKSTALTQHRGKPIELS 1140  
DB 1138 SERIALGPKGSPVTEKEVLTCILCOBEQVKIENNAMVLSACVOKSTALTQHRGKPIELS 1140  
QY 1198 PLCKSLCNTVPIPIPIPOKINSADALAQLLTLARWITQTVLARIISGNIRHAKGNP- 1256  
DB 1201 PLCKSLCNTVPIPIPIPOKINSADALAQLLTLARWITQTVLARIISGNIRHAKGNP- 1256  
QY 1257 IPIFFNQMGDSTLEPHSILSFQVSSIKVNSIKEMVILFATTIYRIGLKVPPDPR 1316

DB 1261 VPVLFNQMGDSTLEPHSILSFQVSSIKVNSIKEMVILFATTIYRIGLKVPPDPR 1320  
QY 1317 VPMLTNSCTCAPTIOAIENLJGDECKPLFGALQNRHNGLKALMQFAVQRIITCQVLIQK 1376  
DB 1321 VPMTWTSCAPTIOAIENLJGDECKPLFGALQNRHNGLKALMQFAVQRIITCQVLIQK 1380  
QY 1377 HLVRLLSVLPNPKSREPTCLLSIDLPHVLVGAFLAPPSLYWDDPVDLQPSVSSSYNHL 1436  
DB 1381 HLARLLSVLPNPKSREPTCLLSIDLPHVLVGAFLAPPSLYWDDPVDLQPSVSSSYNHL 1440  
QY 1437 YLPHLITMAHMLQILLTVDT---GLPLAQVOQEDSEBAHSASSPFAEISQVTSIGCDI 1492  
DB 1441 YLPHLITMAHMLQILLTVDT---GLPLAQVOQEDSEBAHSASSPFAEISQVTSIGCDI 1500  
QY 1493 PGWLVSLKNGITPYLRCAALPHYLICVTPPELHTNSAEGYSALCSVLSLPTNLFL 1552  
DB 1501 PGWLVSLKNGITPYLRCAALPHYLICVTPPELHTNSAEGYSALCSVLSLPTNLFL 1560  
QY 1553 LFQYVDTVRPLIQRCADPALLNCLKQNTVVPYPRKRNLSLIELPDDYSCLLMQASHFR 1612  
DB 1561 LFQYVDTVRPLIQRCADPALLNCLKQNTVVPYPRKRNLSLIELPDDYSCLLMQASHFR 1620  
QY 1613 CPRSADDERKHPVLCFLCGAILCSQNICCOBIVNGEVEGACIFHALHCGAGVCIFLKIRE 1672  
DB 1621 CPRSADDERKHPVLCFLCGAILCSQNICCOBIVNGEVEGACIFHALHCGAGVCIFLKIRE 1680  
QY 1673 CRVVLVEGKARGCAYPAPYLDYGETDPLKRGKGNPLHLSRERYKHLVWQHCIIIEBIA 1732  
DB 1681 CRVVLVEGKARGCAYPAPYLDYGETDPLKRGKGNPLHLSRERYKHLVWQHCIIIEBIA 1740  
QY 1733 RSQETNQMLPGFNQOLL 1749  
DB 1741 RSQETNQMLPGFNQOLL 1757

## RESULT 8

ADJ95455  
ID ADJ95455 standard; protein; 1757 AA.  
XX  
AC ADJ95455;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Mouse Ubiquitin ligase E3alpha I.  
XX  
KW Mouse; enzyme; ubiquitin ligase; E3alpha I; ubiquitin-proteasome pathway;  
KW gene therapy; vaccine; muscular atrophy; cachexia; catabolic disorders;  
KW cancer cachexia; renal cachexia; inflammatory cachexia;  
KW muscle wasting disorder; metabolic acidosis; uremia; burn;  
KW hyperthyroidism; Cushing's syndrome; fasting; denervation atrophy;  
KW diabetes mellitus; sepsis; AIDS wasting syndrome.  
XX  
OS Mus musculus.  
XX  
PN US6706505-B1.  
XX  
PD 16-MAR-2004.  
XX  
PF 28-NOV-2000; 2000US-00724126.  
XX  
PR 08-MAR-2000; 2000US-0187911P.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Han H, Kwak K;  
XX  
DR WPI; 2004-236723/22.  
XX  
XX New nucleic acid molecule, useful for preparing a composition for  
XX diagnosing, treating or preventing diseases associated with human  
XX E3approximately polypeptide, e.g., muscle atrophy.

PS Example 1; SEQ ID NO 15; 104pp; English.

XX The invention relates to a new isolated nucleic acid molecule appearing  
CC as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I  
CC protein appearing as ADJ95442. Also included are a vector comprising the  
CC nucleic acid, a host cell comprising the vector, a process of producing a  
CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the  
CC nucleic acid molecule, a reagent comprising a detectably labelled  
CC nucleotide, and a method for determining the presence of a human E3alpha  
CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid  
CC molecule is useful for preparing a composition for diagnosing, treating  
CC or preventing diseases associated with human E3alpha I polypeptide, e.g.  
CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal  
CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,  
CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting  
CC syndromes. The present sequence represents mouse E3alpha I.

XX Sequence 1757 AA;

Query Match 93.4%; Score 8617; DB 8; Length 1757;

Best Local Similarity 92.1%; Pred. No. 0;

Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;

QY 1 MADEZAGGTERMEISAEPLQTPORLASWDDQVDFYFALHQAOLVPEIYFAEMDPDLE 60  
DB 1 MADEEMDGAERMDVSPPEPLAPORPASWDDQVDFYFALHQAOLVPEIYFAEMDPDLE 60  
QY 61 KOEESVQMSIPTLEWYLFGEHPDICELEKHSAGFOLCGRVFKSGGETTYSRCDCAIDPT 120  
DB 61 KOEESVQMSILPLEWYLFGEHPDICELEKHSAGFOLCGRVFKSGGETTYSRCDCAIDPT 120  
QY 121 CVLMDPCFODSVHKNRYKHHTSTGGPCDCGDEAWKTPGFCVNHPEGRAGTICKENSRC 180  
DB 121 CVLMDPCFODSVHKNRYKHHTSTGGPCDCGDEAWKTPGFCVNHPEGRAGTICKENSRC 180  
QY 181 PLNEEVIQVARKIPSVIKYVEMTIWEKEKLPPELOIREKNERYCYLVNDEHSHVDH 240  
DB 181 PLNEEVIQVARKIPSVIKYVEMTIWEKEKLPPELOIREKNERYCYLVNDEHSHVDH 240  
QY 241 VIYSIQRALDCBLAEQAHLHTTAIDKEGRAVAKAGAYACORAKEDIKHSENVSOHPLHV 300  
DB 241 VIYSIQRALDCBLAEQAHLHTTAIDKEGRAVAKAGAYATCORAKEDIKHSENVSOHPLHV 300  
QY 301 EYLHSEIMAHOKFALRLGSMWNKIMSYSSDFRQIFCOACLEPDSNPCLISRLMLMDA 360  
DB 301 EYLHSEIMAHOKFALRLGSMWNKIMSYSSDFRQIFCOACLEPDSNPCLISRLMLMDA 360  
QY 361 KLYGARKILHLBILFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSVDQ 420  
DB 361 KLYGARKILHLBILFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSVDQ 420  
QY 421 MPTVPTLARHLLEQNVSIVITETLEVLPEYLDNRNKNFPGYSODKLGRIYAVICDLK 480  
DB 421 MPTVPTLARHLLEQNVSIVITETLEVLPEYLDNRNKNFPGYSODKLGRIYAVICDLK 480  
QY 481 YILISKPTITWBLRMQFLEGRFSFLKILTCMQGMEIRROVQGHIEVDPDWEAAIAIQM 540  
DB 481 YILISKPTITWBLRMQFLEGRFSFLKILTCMQGMEIRROVQGHIEVDPDWEAAIAIQM 540  
QY 541 QLNKILLMPQEWACDEBELLVAYKECHAVNRCSTPISSTKTVVQSGHSLKTSYRV 600  
DB 541 QLNKILLMPQEWACDEBELLVAYKECHAVNRCSTPISSTKTVVQSGHSLKTSYRV 600  
QY 601 SEDLVSIHLPLSRTAGHLVLSRLGAVSRLEHFVSPDQVLEVLVEPLRCLVLVAQV 660  
DB 601 SEDLVSIHLPLSRTAGHLVLSRLGAVSRLEHFVSPDQVLEVLVEPLRCLVLVAQV 660  
QY 661 AEMWRNGLSLISQVFFYQDVKCREMYDKOIMLQIGASLMDPNKFLLLVLQRYELAE 720  
DB 661 AEMWRNGLSLISQVFFYQDVKCREMYDKOIMLQIGASLMDPNKFLLLVLQRYELAE 720  
QY 721 FNKTISTKDQDLIKQNTLIEBMLQVLIYVGERVYVPGVGNVTKEBVTREIHLCLIEP 780

RESULT 9

DB 721 FNKTISTKDQDLIKQNTLIEBMLQVLIYVGERVYVPGVGNVTREIHLCLIEP 780  
QY 781 MPHSAIAKLPENENNETGLENNINKVATFKKPGVSGHGVYELKDESLEKDFMYFYHYSK 840  
DB 781 MPHSAIAKLPENENNETGLENNINKVATFKKPGVSGHGVYELKDESLEKDFMYFYHYSK 840  
QY 841 TOHSAEAMQKRRKQENKDBALPPPPPPPCFAPSKVINLNLNCIDMMYILATVFERAID 900  
DB 841 TOHSAEAMQKRRKQENKDBALPPPPPPPCFAPSKVINLNLNCIDMMYILATVFERAID 900  
QY 901 TDSNLWTEGMLQMAFHIALGILREKQOLQKAPREEVTFDPYHKASRLGSSAM---NQM 957  
DB 901 TDSNLWTEGMLQMAFHIALGILREKQOLQKAPREEVTFDPYHKASRLGSSAMNQM 957  
QY 958 LLEKLGIPQLEGQKDMITWLOMPDTVKRLREKSLVATTSGSESIKNDBITHDKKA 1017  
DB 958 LLEKLGIPQLEGQKDMITWLOMPDTVKRLREKSLVATTSGSESIKNDBITHDKKA 1017  
QY 1018 ERKKAEEAARLHROKIMQMSALQKNIETHTKLYDNTSEMPGKEDSIMEBESTPAVS 1077  
DB 1018 ERKKAEEAARLHROKIMQMSALQKNIETHTKLYDNTSEMPGKEDSIMEBESTPAVS 1077  
QY 1078 SERIALGKRGSPVTEKEVLTCLCOEEOVKIENNAMVLSACVKSTALTOHRGKPIBLS 1137  
DB 1078 SERIALGKRGSPVTEKEVLTCLCOEEOVKIENNAMVLSACVKSTALTOHRGKPIBLS 1137  
QY 1138 GEALDPLFMDPLDLAGTYTSGCHVMHACVQKYPFAVQLSSQQRHIVDLFDLESSEYLIC 1197  
DB 1138 GEALDPLFMDPLDLAGTYTSGCHVMHACVQKYPFAVQLSSQQRHIVDLFDLESSEYLIC 1197  
QY 1201 PLCKSLCNTVPIPILOPKINSADALAOULLARMIQTVLARIQVIRHAKGNP- 1256  
DB 1201 PLCKSLCNTVPIPILOPKINSADALAOULLARMIQTVLARIQVIRHAKGNP- 1256  
QY 1257 IPIFNQMGSGSTLEPHSILSPGVESSTKYSNSIKEMVILFATTYIRGLKVPDPDR 1316  
DB 1257 IPIFNQMGSGSTLEPHSILSPGVESSTKYSNSIKEMVILFATTYIRGLKVPDPDR 1316  
QY 1317 VPMLTWTSFCAPTIOAIENLGLDEGKPLFGALQNRHGLKALMQPFAVQRTICPOVLQK 1376  
DB 1317 VPMLTWTSFCAPTIOAIENLGLDEGKPLFGALQNRHGLKALMQPFAVQRTICPOVLQK 1376  
QY 1377 HLVRLLSVLPIKSEDPCLLSIDLPHVLVGAVALPSPSLYWDVDPDIQPSVSSSYNHL 1436  
DB 1377 HLVRLLSVLPIKSEDPCLLSIDLPHVLVGAVALPSPSLYWDVDPDIQPSVSSSYNHL 1436  
QY 1437 YLPHLITWAMLOILLITVDT---GLPLAQVQEDSEEAHSAASSPFAEISQVTSIGCDI 1492  
DB 1437 YLPHLITWAMLOILLITVDT---GLPLAQVQEDSEEAHSAASSPFAEISQVTSIGCDI 1492  
QY 1493 PGWYLVSLKNGITPYLRCAALPFHYLIGVTPPBELHTNSABEGYSALCSYLSLPTNLP 1552  
DB 1493 PGWYLVSLKNGITPYLRCAALPFHYLIGVTPPBELHTNSABEGYSALCSYLSLPTNLP 1552  
QY 1553 LFOEYDVTVPRLQRCWADPALNCLKOKNTVVPYPRKNSLIELPDYSCLLNQASHFR 1612  
DB 1553 LFOEYDVTVPRLQRCWADPALNCLKOKNTVVPYPRKNSLIELPDYSCLLNQASHFR 1612  
QY 1613 CPRSADDERKHPVLCFCGAILCSQNICCOBIIVNGEVEGACIFHALHCGAGVCIPLKIRE 1672  
DB 1613 CPRSADDERKHPVLCFCGAILCSQNICCOBIIVNGEVEGACIFHALHCGAGVCIPLKIRE 1672  
QY 1673 CRVVLVEKARGCAVPAPYLDYGETDPLGKRGNPLHLRERYRKLHLVWQOHCIIIEBIA 1732  
DB 1673 CRVVLVEKARGCAVPAPYLDYGETDPLGKRGNPLHLRERYRKLHLVWQOHCIIIEBIA 1732  
QY 1733 RSQETNQMLFGFNQOLL 1749  
DB 1733 RSQETNQMLFGFNQOLL 1749  
QY 1741 RSQETNQMLFGFNQOLL 1757  
DB 1741 RSQETNQMLFGFNQOLL 1757



ADS86877

ID ADS86877 standard; protein; 1757 AA.

XX

AC ADS86877;

XX

DT 16-DEC-2004 (first entry)

XX

DE Murine E3alpha ubiquitin ligase, muE3I protein sequence.

XX

KW Mouse; E3alpha ubiquitin ligase; muE3I; ubiquitin-proteasome pathway; rapid muscle wasting; fasting; metabolic acidosis; muscle degeneration; kidney failure; renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome; cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome; inflammatory cachexia; hyperthyroidism; denervation atrophy; protein/tissue wasting; energy-protein malnutrition; muscle atrophy; gene therapy; enzyme.

XX

OS Mus musculus.

XX

PN US2004185037-A1.

XX

PD 23-SEP-2004.

XX

XX 15-JAN-2004; 2004US-00758672.

XX

PR 08-MAR-2000; 2000US-0187911P.

PR

PR 28-NOV-2000; 2000US-00724126.

XX

PA (HANH/) HAN H.

PA

PI (KWAK/) KWAK K.

XX

PI Han H, Kwak K;

XX

XX WPI; 2004-707854/69.

DR

XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule useful for treating and/or preventing renal cachexia or inflammatory cachexia.

XX

PS Example 1; SEQ ID NO 15; 115pp; English.

XX

The present invention relates to new orthologue of human E3alpha ubiquitin ligase, huE3alpha1 and huE3alpha11. Most intracellular proteins are degraded through the ubiquitin-proteasome pathway. Proteins are marked for proteasomal degradation by conjugation of ubiquitin to the protein. Conjugation of the ubiquitin molecule involves the activation by E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier, and then interacts with a specific E3 ligase family member. E3 ligase binds to proteins targeted for degradation and catalyses the transfer of ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase determines the specificity of the system. The E3alpha family is the main family of intracellular ligases and is involved in the N-end rule pathway of protein degradation. E3alpha enzyme binds directly to the primary destabilising N-terminal amino acid and catalyses ubiquitin conjugation thereby targeting the protein for degradation. The human E3alpha gene is located on chromosome 15 q. Increased proteolysis through the ubiquitin-proteasome pathway has been determined to be a major cause of rapid muscle wasting including, fasting, metabolic acidosis, muscle degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus, sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen balance, burns, Cushing's syndrome, inflammatory cachexia, hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-protein malnutrition. E3alpha plays a role in the overall increase in ubiquitination that is associated with and may mediate muscle atrophy in catabolic and other disease states. Treatment may be administered by gene therapy, cell therapy and antisense therapy methods. The present sequence is murine E3alpha ubiquitin ligase, muE3I protein sequence.

XX

SQ Sequence 1757 AA;

Query Match

Best Local Similarity 93.4%; Score 8617; DB 8; Length 1757;

Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;

QY 1 MADEBAGTERMEISABELPOTFORLASWDOQVDFTYAFHLHLAQLVPEIFYAEMDPDLE 60  
DB 1 MADEMDGAERMDVSPFPPLAPQRPASWDOQVDFTYAFHLHLAQLVPEIFYAEMDPDLE 60  
QY 61 KOESVQMSIPTPLEWYLFGBDPICLEKLKHSAPOLCGRVKPSGETTYSRCDAIDPT 120  
DB 61 KOESVQMSILTPLEWYLFGBDPICLEKLKHSAPOLCGRVKPSGETTYSRCDAIDPT 120  
QY 121 CVLCHDCFQDSVHKNRYKMTSTGGGFCDCGDTFANKTGPVCYNHBRGAGTIKENSRC 180  
DB 121 CVLCHDCFQDSVHKNRYKMTSTGGGFCDCGDTFANKTGPVCYNHBRGAGTIKENSRC 180  
QY 181 PLNEBVTVOARKIPPSVIVKYVEMTWEEKEKELPELQIREKNERYICVLFNDEHSHYDH 240  
DB 181 PLNEBVTVOARKIPPSVIVKYVEMTWEEKEKELPELQIREKNERYICVLFNDEHSHYDH 240  
QY 241 VIYSLQALDCELAQAQLHTTAIDKEGRRAVAGAYAACQBAKEDIKSHSNVSOHPLHV 300  
DB 241 VIYSLQALDCELAQAQLHTTAIDKEGRRAVAGAYAACQBAKEDIKSHSNVSOHPLHV 300  
QY 301 EVLHSEIWAHQKPAIRLGSNNKIMSYSSDPRQIFCOACLEEDPDSENPCLISRLMLWA 360  
DB 301 EVLHSEIWAHQKPAIRLGSNNKIMSYSSDPRQIFCOACLEEDPDSENPCLISRLMLWA 360  
QY 361 KLYGARKILHLLIFSSFFMEMEYKLFAMEFVYKQLOKEYISDDHRSISITALSQ 420  
DB 361 KLYGARKILHLLIFSSFFMEMEYKLFAMEFVYKQLOKEYISDDHRSISITALSQ 420  
QY 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNNKNFQGYSDQKLGKVAVICDLK 480  
DB 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNNKNFQGYSDQKLGKVAVICDLK 480  
QY 481 YILISKPTIWTERRMQFLEGFRSLKILTCMOGHEIRROVGOHIEVDPDWEAAIAIQM 540  
DB 481 YILISKPTIWTERRMQFLEGFRSLKILTCMOGHEIRROVGOHIEVDPDWEAAIAIQM 540  
QY 541 QKNIILLMFQEWACADELLVAVKECHKAVMRCTSTFSSKTVVSCGHSLETYSYV 600  
DB 541 QKNIILLMFQEWACADELLVAVKECHKAVMRCTSTFSSKTVVSCGHSLETYSYV 600  
QY 601 SEDLVSIHLPLSRTLAGLHVLRLSRLGAVSRHLHFVSPEDQVFLVEYPLCLVLVAQV 660  
DB 601 SEDLVSIHLPLSRTLAGLHVLRLSRLGAVSRHLHFVSPEDQVFLVEYPLCLVLVAQV 660  
QY 661 AEMWRRNGLSISQVFFYQDVCKREEMVDKDIIMLOIGASLMDPNKLLVLQRYELAEA 720  
DB 661 AEMWRRNGLSISQVFFYQDVCKREEMVDKDIIMLOIGASLMDPNKLLVLQRYELAEA 720  
QY 721 FNKTIKTDQDLIKQYNTLIEBQNVISVITETLLEVLPEYLDNNKNFQGYSDQKLGKVAVICDLK 780  
DB 721 FNKTIKTDQDLIKQYNTLIEBQNVISVITETLLEVLPEYLDNNKNFQGYSDQKLGKVAVICDLK 780  
QY 781 MPHSAIAKPLPENENNETGLNVLNKVATPKKPGVSGHGVYELKDESUKDNFMFYHYSK 840  
DB 781 MPHSAIAKPLPENENNETGLNVLNKVATPKKPGVSGHGVYELKDESUKDNFMFYHYSK 840  
QY 841 TQSKAEHMQKRRKQENKDALPPPPPEPCFAPSKVINLLCNDIMMYILRTVPERAID 900  
DB 841 TQSKAEHMQKRRKQENKDALPPPPPEPCFAPSKVINLLCNDIMMYILRTVPERAID 900  
QY 901 TDSNLWTBGLMQLAFHILALGLLEBQKQOLKAPBEEVTFDFYHKASRLGSSAM---NTQM 957  
DB 901 TESNLWTBGLMQLAFHILALGLLEBQKQOLKAPBEEVTFDFYHKASRLGSSAMNTQM 957  
QY 958 LLEKLGKIPQLEGQKDMITWILQMPDVTVKRLREKSLIVATTSSESINKDEITHDKSKA 1017  
DB 958 LLEKLGKIPQLEGQKDMITWILQMPDVTVKRLREKSLIVATTSSESINKDEITHDKSKA 1017  
QY 1018 ERKKAARLHROKIMQMSALQKNFLETHKLMYDNTSEMPGKEDSTMEESTPVSVDY 1077  
DB 1021 ERKKAARLHROKIMQMSALQKNFLETHKLMYDNTSEMPGKEDSTMEESTPVSVDY 1077

```
QY 1078 SRALGPKEGSPVTEKEVLTCLCOBEQVKLENNAMVLSACVOKSTALTQHRGKPIELS 1137
DB 1081 SRALGPKEGPAVTEKEVLTCLCOBEQVKLENNAMVLSACVOKSTALTQHRGKPVDFHL 1140
QY 1138 GEALDPLFMDPOLAYGTYTGSCHVNHAVCWQKYFEAVQLSSQORIHVDLFDLESGEYLC 1197
DB 1141 GETLDPLFMDPLAHTGTYTGSCHVNHAVCWQKYFEAVQLSSQORIHVDLFDLESGEYLC 1200
QY 1198 PLCKSLCNTVPIPILOPKINSNADALAQLLTLARWTQTVLARISGVNIRHAKGNP- 1256
DB 1201 PLCKSLCNTVPIPILOPKINSNADALAQLLTLARWTQTVLARISGVNIRHAKGEADA 1260
QY 1257 IPIFNGQMGDSFLPHSILSCGVSSIKYSIKEMVILFATTIYRIGLKVPDPDERDR 1316
DB 1261 VPVLFNGQMGDSFEPHLSILSGVQSSVKSISIKEMVILFATTIYRIGLKVPDPDELDR 1320
QY 1317 VPMLTWSTCAFTIQTALENLLGDEKPLFGALQNRHNGIKALMQFAVQRIICPOVLIQK 1376
DB 1321 VPMWTWSTCAFTIQTALENLLGDEKPLFGALQNRHSGIKALMQFAVQRAICPOVLIHK 1380
QY 1377 HLVRLSVVLPNIKSEDTPCLLSIDLPHVLGAVLAPPVSLYWDVDPVQLQPSVSSYNHL 1436
DB 1381 HVARLLSVLPNLQSENTPGLLSVDLPHVLGAVLAPPVSLYWDVTVDLQPSPLSSYNHL 1440
QY 1437 YLPHLITWAHMLQILLTVDT----GLPLAQVOEDSEEAHSAPFAEISQYTSIGCDI 1492
DB 1441 YLPHLITWAHMLQILLTVDTDLSGPPLEAGBEDSEEAFCASAFVFSQHTDGLTGCGA 1500
QY 1493 PGWYLVSVLKNGTTPYRCAALPFHYLLGVTPPEELHTNSARGESYALCSYLSLPTNLPL 1552
DB 1501 PGWYLVSVLRNGITPYRCAALPFHYLLGVTPPEELPANSARGESYALCSYLSLPTNLPL 1560
QY 1553 LFQEWYDVTVRPLQRCWADPALLCNLCQKNTVYRPRKNSIELPDDYSCILNQASHFR 1612
DB 1561 LFQEWYDVTIRPLQRCWGPALLKSLQKSAVVYRPRKNSIELPDDYSCILNQASHFR 1620
QY 1613 CPRSADDERKHPVLCFCGAILCSQNICCOEIVNGEVEGACIFHALHCGAGVCIFFKIRE 1672
DB 1621 CPRSADDERKHPVLCFCGAILCSQNICCOEIVNGEVEGACVFFHALHCGAGVCIFFKIRE 1680
QY 1673 CRVVLVEGARGCAYPAPVLDYGETDPCGKRGNPLHLRSRERYKHLVWQOHCIIERTA 1732
DB 1681 CRVVLVEGARGCAYPAPVLDYGETDPCGKRGNPLHLRSRERYKHLVWQOHCIIERTA 1740
QY 1733 RSQETNQMLFGFNWQLL 1749
DB 1741 RSQETNQMLFGFNWQLL 1757

RESULT 10
ADI16325
ID ADI16325 standard; protein; 908 AA.
XX
AC ADI16325;
XX
DT 22-APR-2004 (first entry)
XX
DE Human protein modification and maintenance molecule (PMMW) protein #10.
XX
KW human; protein modification and maintenance molecule; PMMW;
KW gastrointestinal disorders; peptic ulcer; Crohn's disease;
KW cardiovascular disorders; hypertension; congenital heart disease;
KW autoimmune disease; inflammatory disease; AIDS; anaemia;
KW developmental disorder; Cushing's syndrome; tubular acidosis;
KW epithelial disorder; eczema; scabies; neurological disorder;
KW Alzheimer's disease; multiple sclerosis; infection; cancer.
XX
OS Homo sapiens.
XX
PN WO2003100016-A2.
XX
XX 04-DEC-2003.
XX
PD
```

```
PP 22-MAY-2003; 2003WO-US016498.
XX
PR 22-MAY-2002; 2002US-0383491P.
PR 24-JUN-2002; 2002US-0391378P.
PR 22-JUL-2002; 2002US-0397921P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Khare R, Bulloch SA, Swarnakar A, Elliott VS, Marquis JP,
PI Mason PM, Chawla NK, Ramkumar J, Kable AE, Hafalia AJA, Lee SY,
PI Tran UK, Yue H, Becha SD, Griffin JA, Chang H, Jiang X, Jackson AA,
PI Richardson TW, Lal PG, Yao MG, Lu Y, Warren BA, Jin P, Wilson AD,
PI Gietzen KJ,
XX
DR WPI; 2004-035124/03.
DR N-PSDB; ADI16377.
XX
PT New protein modification and maintenance molecules, useful for diagnosing
PT or treating e.g. peptic ulcer, hypertension, rheumatic fever, AIDS,
PT Cushing's syndrome, Alzheimer's disease, multiple sclerosis, stroke or
PT cancers.
XX
PS Claim 1; SEQ ID NO 10; 419pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC protein modification and maintenance molecules (PMMW). The DNA and
CC protein sequences of the invention are useful for the diagnosis and
CC treatment of disorders associated with expression of PMMW, such as:
CC gastrointestinal disorders (e.g. peptic ulcer and Crohn's disease),
CC cardiovascular disorders (e.g. hypertension and congenital heart
CC disease), autoimmune or inflammatory disease (e.g. AIDS and anaemia),
CC developmental disorders (e.g. Cushing's syndrome and tubular acidosis),
CC epithelial disorders (e.g. eczema and scabies), neurological disorders
CC (e.g. Alzheimer's disease and multiple sclerosis), infections and cancer.
CC The present amino acid sequence represents a human PMMW protein of the
CC invention.
XX
SQ Sequence 908 AA;
Query Match 51.7%; Score 4768; DB 8; Length 908;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 902; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 12 MEISAEPLPOTQRLASWMDQVDFYTAFLHQAQLVPEIYFAEMDPDLEKQESVQMSIF 71
DB 1 MEISAEPLPOTQRLASWMDQVDFYTAFLHQAQLVPEIYFAEMDPDLEKQESVQMSIF 60
QY 72 TPLEWYLFGEHPDICLEKLKHSAPOLCGRVFKSGETTYSRCDCAIDPTCVLCMDCFQDS 131
DB 61 TPLEWYLFGEHPDICLEKLKHSAPOLCGRVFKSGETTYSRCDCAIDPTCVLCMDCFQDS 120
QY 132 VHKHRYKMHHTSTGGPCDCGDTBAWKTGPPCVNHEPGRAGTIKENSRCPLNEEVIVQAR 191
DB 121 VHKHRYKMHHTSTGGPCDCGDTBAWKTGPPCVNHEPGRAGTIKENSRCPLNEEVIVQAR 180
QY 192 KIPFSVIKYVEMTIVBEEKELPPELQIREKNRYCYLVFNDEHSHSDHYVLSQALDC 251
DB 191 KIPFSVIKYVEMTIVBEEKELPPELQIREKNRYCYLVFNDEHSHSDHYVLSQALDC 240
QY 252 ELAEALHTTAIDKEGRANVAGAYACQAEKEDIKSHSENVVSHPLHVEVLHSHMAHQ 311
DB 241 ELAEALHTTAIDKEGRANVAGAYACQAEKEDIKSHSENVVSHPLHVEVLHSHMAHQ 300
QY 312 KPALRLGSMWNKIMSYSSDPKQIFCOACLRREPDPSENPCILSRMLWDAKLYKGARKILH 371
DB 301 KPALRLGSMWNKIMSYSSDPKQIFCOACLRREPDPSENPCILSRMLWDAKLYKGARKILH 360
QY 372 ELIFSSPFMEYKYLKLFAMBFVKYKQLQKEYISDDHDSISITALSVMQFTVPTLARHL 431
DB 361 ELIFSSPFMEYKYLKLFAMBFVKYKQLQKEYISDDHDSISITALSVMQFTVPTLARHL 420
QY 432 IEQNVSIVITETLLEVLPEYLDNRNKNFPGYSODKLGRTVAVTCDLKYLISKPTTWT 491
```

Db 421 IBEQNVISVITTELLVLEPEYLDNRNKNFQGSQDKLGRVAVICDLKYLISKPTIWT 480

Qy 492 ERLRQFLGFSFLKILTCMQGMEIRROVGOHLEVPDWEAAIAIOMQLNILLMFOE 551

Db 481 ERLRQFLGFSFLKILTCMQGMEIRROVGOHLEVPDWEAAIAIOMQLNILLMFOE 540

Qy 552 WCACDEELLVAYKECHKAVMRCSTSFSSSTVTVQSCGHSLETYSYRVEDLSIHLPL 611

Db 541 WCACDEELLVAYKECHKAVMRCSTSFSSSTVTVQSCGHSLETYSYRVEDLSIHLPL 600

Qy 612 SRTLAGLHVLRLGAVSRHLEFVDFEDQVFLVEYPLRLCLVLVAQVVAEMRRNGLSL 671

Db 601 SRTLAGLHVLRLGAVSRHLEFVDFEDQVFLVEYPLRLCLVLVAQVVAEMRRNGLSL 660

Qy 672 ISQVFFYQDVKCREEMDKIIMLQIGASLMDPNKFLILVLQRYELAZAFNKTIKDOOD 731

Db 661 ISQVFFYQDVKCREEMDKIIMLQIGASLMDPNKFLILVLQRYELAZAFNKTIKDOOD 720

Qy 732 LIKQYNTLIEEMQLVLIYVGRYVPGVGNVTKKEVTWREIITHLLCIEPMHSAIAKNLP 791

Db 721 LIKQYNTLIEEMQLVLIYVGRYVPGVGNVTKKEVTWREIITHLLCIEPMHSAIAKNLP 780

Qy 792 ENNETGLNVINKVATPKPGVGHGYELKDSLKDFNMFYHYSTQHSKAEHQK 851

Db 781 ENNETGLNVINKVATPKPGVGHGYELKDSLKDFNMFYHYSTQHSKAEHQK 840

Qy 852 KRRQENKDEALPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 911

Db 841 KRRQENKDEALPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 900

Qy 912 QMAFHI 917

Db 901 QMDWEV 906

RESULT 11

ADJ95444

ID ADJ95444 standard; protein; 1755 AA.

AC ADJ95444;

XX

03-JUN-2004 (first entry)

XX

Human Ubiquitin ligase E3alpha II.

XX

Human; enzyme; ubiquitin ligase; E3alpha II;

KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;

KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;

KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;

KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;

KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome.

XX

Homo sapiens.

OS

XX

US6706505-B1.

PN

XX

16-MAR-2004.

PD

XX

28-NOV-2000; 2000US-00724126.

PF

XX

08-MAR-2000; 2000US-0187911P.

PR

XX

(AMGE-) AMGEN INC.

PA

XX

Han H, Kwak K;

PI

XX

WPI: 2004-236723/22.

DR

XX

N-PSDB; ADJ95443.

DR

XX

New nucleic acid molecule, useful for preparing a composition for

PT diagnosing, treating or preventing diseases associated with human

PT E3approximate polypeptide, e.g., muscle atrophy.

XX

PS Example 2; SEQ ID NO 4; 104pp; English.

XX

The invention relates to a new isolated nucleic acid molecule appearing

CC as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I

CC protein appearing as ADJ95442. Also included are a vector comprising the

CC nucleic acid, a host cell comprising the vector, a process of producing a

CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the

CC nucleic acid molecule, a reagent comprising a detectably labelled

CC nucleotide, and a method for determining the presence of a human E3alpha

CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid

CC molecule is useful for preparing a composition for diagnosing, treating

CC or preventing diseases associated with human E3alpha I polypeptide, e.g.

CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal

CC cachexia, inflammatory cachexia, muscle wasting disorders associated with

CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,

CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting

CC syndrome. The present sequence represents human E3alpha II.

XX

Sequence 1755 AA;

Query Match 46.9%; Score 4327; DB 8; Length 1755;

Best Local Similarity 47.2%; Pred. No. 0;

Matches 841; Conservative 339; Mismatches 543; Indels 58; Gaps 27;

Qy 1 MADE---EAGTER--MEISAEPLQTPQRLASWMDQVDFTAFLLHLAQLVPIYFAEM 55

Db 1 MASELEPEVQAIKSLLECSAE-----EIAAGKWLQATDLTREVVQHLAVVPKIYCRGP 54

Qy 56 DPDLKESQESVQMIPTPLEWYLFGEPPDICLEKLKHSQ-AFQLGGRVFKSGEYTSYCRD 114

Db 55 NPPFKEDMLAQHVLGLGPMWYLGQEDPAFPFKLEQANKPSHLGCRVFKYGEYTSYCRD 114

Qy 115 CAIDPTCVLNDQFQDSVHKHRYKMTSTGGGFCDCGDTAEWTKGPCVNVHPEGRAG-T 173

Db 115 CAVDPTCVLMCEFLGSIHRDHYRMTTSGGGGFCDCGDTAEWKEGPGYCKQHELTSTIE 174

Qy 174 IKENSRCPLNEBEVIVQARKIPPSVIKYVVENTIWEKEKELPPELQIKENRYTCVLFND 233

Db 175 EEDPLVHLSDEVARTYNIPTFRVAVEILTWEKESLPALEWKESTYTCVLFND 234

Qy 234 EHSYDVHYISLQALDCELAELAHHTTAIDKEGRRAVAGAYACQAKSIDIKSHSENV 293

Db 235 EVHTYQVIYTLQKAVNCTQKEAIGFATTVDRGRSSVRYGDFQYCEQAKSVIRNTSRQ 294

Qy 294 SQHPLHVEVLHSETMAHQKFAIRLGSWMNKIWSYSSDPQIFCQACLEEPDSENPCIS 353

Db 295 TK-PLKQVHMSSIVAHQNFGLKLLSLWGSIIIGYSDGLRRLICQVLQEGDGENSLVD 353

Qy 354 RLMLWDKLYKGARKILHELIFSSFFMEMEYKCLFAMEFVKYKQLQKEYISDDHDSIS 413

Db 354 RLMLSDSKLWKGARSVYHQLFWSSLLMDLYKLLFAVFAKNYQQLQDFMEDDHERAVS 413

Qy 414 ITALSVMQFTVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNKNFQGS---QDKLG 470

Db 414 VTALSVPQFTAPTARMLITBENLMSIIKTFFMDHL-RHRDAQGRFPQERYTALQAFKR 472

Qy 471 RVYAVICDLKYILISKPTIWTTERLMQFGRFLKILTCMQGMEIRROVGOHLEVPD 530

Db 473 RVQSLILDLKYVLIISKPTENSDRLQKFLGDFAPLELLKCMQGMPIRITRQVGOHIEHP 532

Qy 531 DWEAAIAIQMOLKNILLMFOEWACDEELLVAYKECHKAVMRCSTSFSSSTVTVQSC 589

Db 533 EWEAAFTLQMKLTHVISMMQDCASDEKVLIEYKCLAVLMQCHGGYTDGEQPTLSIC 592

Qy 590 GHSLETYSYRVEDLSIHLPLSRTLAGLHVLRLGAVSRHLEFVDFEDQVFLVEYPL 649

Db 593 GHSVETIRYCVSQESVHLPVSRLLAGLHLLSKSEVAYKFPPELPLSELSPMPLIHP 652

Qy 650 LRLCLVLAQVVAEMRRNGLSLISQVFFYQDVKCREEMDKIIMLQIGASLMDPNKFL 709

Db 653 LRLCLVLAQVVAEMRRNGLSLISQVFFYQDVKCREEMDKIIMLQIGASLMDPNKFL 712

Qy 710 LVLQRYELAZAFN-----KTIKTK--DQDLIKQYNTLIEEMQLVLIYVGRYVPGVGNV 762



[illegible]

RESULT 13	
ADJ95446	
ID	ADJ95446 standard; protein; 1755 AA.
XX	
XX	
AC	ADJ95446;
XX	
DT	03-JUN-2004 (first entry)
XX	
XX	
XX	Mouse Ubiquitin ligase E3alpha II.
KW	Mouse; enzyme; ubiquitin ligase; E3alpha II;
KW	ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;
KW	cachexia; catabolic disorder; cancer cachexia; renal cachexia;
KW	inflammatory cachexia; muscle wasting disorder; metabolic acidosis;
KW	uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;
KW	denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome.
OS	Mus sp.
XX	
PN	US6706505-B1.
XX	
PD	16-MAR-2004.



XX PF 28-NOV-2000; 2000US-00724126.  
 XX PR 08-MAR-2000; 2000US-0187911P.  
 XX PA (AMGE-) AMGEN INC.  
 XX PI Han H, Kwak K;  
 XX DR WPI; 2004-236723/22.  
 XX DR N-PSDB; ADJ95445.  
 XX PT New nucleic acid molecule, useful for preparing a composition for  
 PT diagnosing, treating or preventing diseases associated with human  
 PT E3approximately polypeptide, e.g., muscle atrophy.  
 XX PS Example 2; SEQ ID NO 6; 104pp; English.  
 XX CC The invention relates to a new isolated nucleic acid molecule appearing  
 CC as ADJ95441(or its complement) encoding a ubiquitin ligase E3alpha I  
 CC protein appearing as ADJ95442. Also included are a vector comprising the  
 CC nucleic acid, a host cell comprising the vector, a process of producing a  
 CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the  
 CC nucleic acid molecule, a reagent comprising a detectably labelled  
 CC nucleotide, and a method for determining the presence of a human E3alpha  
 CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid  
 CC molecule is useful for preparing a composition for diagnosing, treating  
 CC or preventing diseases associated with human E3alpha I polypeptide, e.g.  
 CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal  
 CC cachexia, inflammatory cachexia, muscle wasting disorders associated with  
 CC metabolic acidosis, uremia, burns, hyperthermoidism, Cushing's syndrome,  
 CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting  
 CC syndrome. The present sequence represents mouse E3alpha II.  
 XX CC Sequence 1755 AA;  
 SQ

Query Match 46.4%; Score 4278; DB 8; Length 1755;  
 Best Local Similarity 46.5%; Pred. No. 0;  
 Matches 835; Conservative 341; Mismatches 534; Indels 84; Gaps 28;

QY 1 MADE---EAGGTER--METSALPQPPQRLASWQDQVFTAFLLHQAOLVPELYFAEM 55  
 DB 1 MASEMPEVQADRLSLECSAE-----EAGRWLQATDNLREYVQHLACHVFKYICRGP 54  
 QY 56 DPDLKQESVQMSIPTLEWYLFGBDDPDICLEKLKHSQ-AFQCGRVFKSGETTSYCRD 114  
 DB 55 NPPQKEDTLAHLILGPNWETICAEDPALGPPKLEQANKPSHLCGRVFKVGEPTYSCRD 114  
 QY 115 CAIDPTCVLCMDQFQDSVHNHRYKMTSTGGGFCDCGDTBAWGTGPPCNNHEPGRAGTI 174  
 DB 115 CAVDPTCVLCMEFLGSIHRDHYRMTTSGGGGFCDCGDTBAWKEGFCYQKHLSSRYV 174  
 QY 175 -KENSRCPLNEEVIQARKIFPSVIKVVYEMTWEEKELPPELQIREKNERYCVLND 233  
 DB 175 EEDPLVLHUSEDVIARTNYFALMFRYAVDIITWEKESLPEDELEVAESDYTYCMLND 234  
 QY 234 EHSYDHVYISLQALDCELABAQLHTTATIDKEGRVAVAGAYAAQAEKBDIKHSENV 293  
 DB 235 EVHTYEQVITYLQKAVNCQKEAIGPATTVDRGRPRVYGFQYCDQAKTIVNTSRQ 294  
 QY 294 SQHPLHVEVLHSEIMAHQKFAIRLGSWMNKIMSYSDFPQIFCQACLREEPSNCPILIS 353  
 DB 295 TK-PLKVVQVHMSVAHQNFGLKALSWLGSYSDGLRILLCQVGLQSGPDGNSSLVD 353  
 QY 354 RLMLWDKLYKARKTLHELIPSSFPNEMEKYKLFAMEVYKYKQLOKEYISDDHDSIS 413  
 DB 354 RLMLNDSKLKAGARSYYHQLFSSSLMDLKYKLFALRFAYRQRLQDFMEDDHERAVS 413  
 QY 414 ITALSVMQMTVPTLARHLEEQNVISVITETLEVLPEYLDNRNKNFQGS---ODKLG 470  
 DB 414 VTALSVOFTAPTALRMLLTENMLTVIIKAFMDHL-KERDAQGRFOFERYTALQAFKR 472  
 QY 471 RYVAVICDLKYLISKPTIWTERRMQFLGFRSFLKILTCMQGMBEIRRQVQGHIEVDP 530

DB 473 RVQSLILDLYVLIISKRPTWSDLELQKFLQGGDAFLELLKCKMQGMDPITRQVQGHIEMP 532  
 QY 531 DWEAAIAIQOMQLNILLFQEWCAQDELLLVAYKECHAVMRCSTSFSSSKTVQVS-C 589  
 DB 533 EWEAAFTLQMLKLVHVSVMQDWCALDEKVLIERAYKKCLAVLTOCHGFTDGRQPTLSIC 592  
 QY 590 GHSLETYSYRVEDLVSHLPLSRILAGLHVLSRLGAVSRHLHFVSPDFQVEVLVEYP 649  
 DB 593 GHSVETIRYCVSQEKSIVHLPISSLRILAGLHVLSKSEVAYKPPPELLSELSPMPLIEHP 652  
 QY 650 LKCLVLAQVAEMRRNGLSLISOVFFYODVKCREMYDKDIIMLOICASIMDNKPTLL 709  
 DB 653 LKCLVLCQVHAGMWRNGFSLVQIYYTHNVKCRREMFDKDIIMLOTVGSMMDNHFLEH 712  
 QY 710 LVLYRYELAEAFN-----KTISTK--DQDLIKQYNTLIEMLQVLIYIVGSRVYVPGVNV 762  
 DB 713 INLSRFELVQLPSTPDYKGRPSSEVTHKDVQNNLTIEEMLYLIIMLVGERFNPQGVQV 772  
 QY 763 TKEEVTMRBIIHLLCIBPMPHSAIANLPENNETGTLENVINKVATFKKPGVSGHVYE 822  
 DB 773 AATDEIKREIITHQLSIKPMASSELVKSLPEDENKETMESVIESVAHFKKPGLTGRGMYE 832  
 QY 823 LKDSLSKDPNMYFYHYSKTQHSKASHMOKRQKQENKDEALPPPPPEPCPAFSKVINLL 882  
 DB 833 LKPECAKGFNLYFYHFSRAEOSKABEORKLKREKEDTALPPALPPCPPLFASLVNLL 892  
 QY 883 NCDIMMYLITVFRRAIDTDSNLTNEGMLQMAFHILALGLLEKQKQLOKAPSEEV-TFDF 941  
 DB 893 QCDVMLYIMTILQWAVEHSGSAMESMLQVHLHIGMALQSEKHLENAVEGHVQTFPFE 952  
 QY 942 YHKASRLGSSANN---IQMLLEKLGIPQLEGQKDMITWLOMPDTVKRLREKSLIVAT 998  
 DB 953 TOKISKPGDAPHNSPSILAMLETQLQWAPSLEAHKDMIRWLLKWFNAIKKIRE--CSSSP 1010  
 QY 999 TSGSESIKNDEITHDKEKAEKRAEAAHLHROKIMAOALSALOKKFIETHKLWYDNTSM 1058  
 DB 1011 VAEAGTTIMBESSRODKAEARKKRAIARLRREKIMAOEMSEMRHPIIDENKELFOOTLEL 1070  
 QY 1059 PGKEDSIMEESTPAVSYSRIALGPKRGPSVTEKEVLTICLQBEQEVKIENNAVLISA 1118  
 DB 1071 DTSASATL--DSSPPVSDAALTALGPAQTVQVPEPQFVTCILQBEQEVTVSGRAVLAA 1128  
 QY 1119 CVQKSTALTQHRKPIELSGEALDPLFMDPLAYGYTTGSCGHVMAVCWKYFRAVQUS 1178  
 DB 1129 FYQSTVLSKORTKI-ADPEKYDPLFMHPDLSCGTHTSCGHVMAHCHQYFDSVQAK 1187  
 QY 1179 SQO-----RIHVDLPDESSEVLCPLCKSLCNVTPIPILOPKINSENAALQALLTIA 1233  
 DB 1188 EORROORLRLHTS-YDVENGEPCLPCLBCLSNVTIPLL-LPPRSILSRRLN-FSDOPDLA 1244  
 QY 1234 RWIQTVLARISGYNIT---RHAKGEN-----PIPIFFNQMGDSTLEFHSILSF 1278  
 DB 1245 QWTRAVTQIKVQVQMLRKHNAADTSSSEDTAMNIIPIPEGFRP-----DFY----- 1292  
 QY 1279 GVSESIKYSNKGWVILPATIYRIGLKVPPDERDPVPMLTWTCAFTIQAIENLLGD 1338  
 DB 1293 ---PRNPYSDSIKEMLTFTGTAAYKVLKGVHNEGDRPVILCWGTCAVTIQSIEBRLSD 1349  
 QY 1339 EGKPLFGALQNRQHNGKALMQPAVAQRITCQOVLIQKHLVLLSVLNPISSEDTPCLL 1398  
 DB 1350 EKPYPGLPCRLDCLSLRTRFAAAHWTVALLPVVQGHFCKLFASLVPDSYEDLPCLTL 1409  
 QY 1399 STDLPVHLVGAVLAPPSLYWDDPVDLQPSVSSVSNHLYLPHLITMAHMLQILL---TVD 1455  
 DB 1410 DIDMFHLVGLVLAIPALQCCD-----FSGSLATGDHLHIFHLVTMAHIVQILLTSCTEE 1464  
 QY 1456 TGLPLAQVQDESEEAHSASSPAEISQYTSIGCDIPGWLVLWSLKNKITPYLRCALF 1515  
 DB 1465 NGMD--QENPTGEBELAILSLHKLTHQVTSALKEAPSGHWRVSRVAAIMPFLKCSALF 1522  
 QY 1516 FHYLLGVTPPELHTNSAGEYSALCSYLSLPTNIFLLFOEYDWTDTVRPLLRWCADPALL 1575



Db 773 AATDEIKREIHHQSLIKPMASSELVLSLPEDENKTCMESVIESVAHFKKPGLTGRGWE 832  
Qy 823 LKDESILKDFMYPIYHYSKTOHSAKAEHMQKRRQKQENKDRALPPPPPPPCAFPSKVINLL 882  
Db 833 LKPECAKEFNLYTHSRABQSKAEARQKLRKENKEDTALPPALPPFCPLPASLVNII 892  
Qy 883 NCDIMVILITVPERAIDTDSNLWTEGMLQWAFHILALGLERKQOLQKAPESV-TPDF 941  
Db 893 QCDVLMYIMTIIQWAVEHGHGSWSSENLQRLVHLIGMALQOEKHLLENAVEGHVQTFP 952  
Qy 942 YHKASRLGSSAMN---IQMLEKLGKIPQLEGOKDMITWILQMPDVTVKRLREKSCLIIVAT 998  
Db 953 TQKISKPGDAPHNSPSILAMLETIQNAPSLEAHKOMIRWLLKMFNAIKKIRE--CSSSP 1010  
Qy 999 TSGSESINKDEITHDKAKRKKKAKAEARLRHOKIQAQMSALQKQNFETHKLMYDNTSEM 1058  
Db 1011 VAEABGTMEESSRDKDKAKRKKKAEIARLRREKIKQAQMSQEMQHFIDENKELFQOTLEL 1070  
Qy 1059 PGKEDSTMEESSTPAVSDYSRIALGPKRGSVTEKEVLTCLCOEBOEVKIENNAVLISA 1118  
Db 1071 DTSASATL--DSBPPVSDAALTALGPAQTQVPPRPQVPTCILCOEBOEVTVGSRAWVLA 1128  
Qy 1119 CVQKSTALTQHRGKPIELSGEALDPLFMDPLAYGTGSCGHVMAVCHQKFFEAVALS 1178  
Db 1129 FVQSTVLSKORTKI-ADPEKYDPLFMHDPDLSCGTHGSCGHVMAHCHQWIFDSVQAK 1187  
Qy 1179 SQO-----RIHVDLPDESGEYLCPLCKSLCNTVPIPIQPOKINSENADALALTLA 1233  
Db 1188 EQRQORLRHTS-YDVNGEFLCPLCECLSNTPVILL-LPPRSILSRRLN-FSDQPDLA 1244  
Qy 1234 RWIQTVLARISGNYI---RHAKEN-----PIPIFNQMGDSTLEFHSILSF 1278  
Db 1245 QMTRAVTQQIKVQMLRKKNAAATSSSEDETEAMNIIPIEGFRP-----DFY----- 1292  
Qy 1279 GVSEIKYNSIKEMWILPATTYIRIGLVKPPDERPRVPLMTSTCAPTIAIENLGD 1338  
Db 1293 ---PRPYSIISIKEMLTFTGTAAYKGLKVHPNEGDRVPLFCWGTCAITQISRIELSD 1349  
Qy 1339 EGKPLFGALQNRQHGKALKMQFAVAQRITCPOVLQKHLVRLSVLVPNIKESDTPCLL 1398  
Db 1350 EEPVFGPLFCRLDDCLSLTRFAAAHWTVALPVVQGHFCKLPASIVPSDSVEDLPCIL 1409  
Qy 1399 SIDLFVILVGAVALPSPLYWDDPDVLOQPSVSSSYNNHLYLPHLITWAMLMQILL----TVD 1455  
Db 1410 DIDMFHLLVGLVLAFFALQOCD-----FSGSSLATGDLHIFHLVWMAHIVQIILLTSCTE 1464  
Qy 1456 TGLPLAQVQDSBEAHSASSFFAEISQYTSIGSCDIPGNYLVSLKNGITPYLRCAALF 1515  
Db 1465 NGMD--QENPTGEEELAILSLHTLHQYTGSALKBAPSGHMLWRSVRAAIMPFLKCALF 1522  
Qy 1516 FHYLLGVTPPEELHTNSAEGEYALCSYLSLPTNLFLFOEYWDTVRPLLQRCADPALL 1575  
Db 1523 FHYLVGVAPPDQLQV-SGTHPEHLNLYLSLPTNLHLFOENSIMNSLIESQNSVEK 1581  
Qy 1576 NCLQKQNTVVRPRKNSLIELPDDYSCLLNQASHFRCPASADERKHVPLCLFCGAILC 1635  
Db 1582 RYLNGERGATSPRGANKLIDLPEYSSLIQNASNFPCKSGDKSRAPTLCLVCGSLLC 1641  
Qy 1636 SQNICCOEIVNGEVEGACIFALHCGAGVCIELKIRECRVTVVEGARGCAYAPVLDY 1695  
Db 1642 SOSYCCOALEGDEVDGACTAHTYSCSGAGIFLRVRECQVFLAGKTGKCFYFPYDDY 1701  
Qy 1696 GETDPGLKRGNPLHLRERVKHLVWQOHCIIIEIARSQBTNQMLPGFNQOLL 1749  
Db 1702 GETDQGLRGNPLHLCOERFKYQKLMQOHSITEEIGHAQAQTLVGIDWQHL 1755

RESULT 15

AAB93464

ID AAB93464 standard; protein; 811 AA.

XX

AC AAB93464;

XX

DT 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:12732.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000BP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isoqai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-

XX length cDNAs defined in the specification, and for the detection and/or

XX diagnosis of the abnormality of the proteins encoded by the full-length

XX cDNAs.

XX Claim 8; SEQ ID NO 12732; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-

XX length cDNAs defined in the specification. Where a primer set comprises:

XX (a) an oligo-dr primer and an oligonucleotide complementary to the 5602

XX complementary strand of a polynucleotide which comprises one of the 5602

XX nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in the

XX specification. The primer sets can be used in antisense therapy and in

XX gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893

XX represent human amino acid sequences; and AAH13629 to AAH13632 represent

XX oligonucleotides, all of which are used in the exemplification of the

XX present invention

XX Sequence 811 AA;

Query Match 45.6%; Score 4209; DB 4; Length 811;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 811; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 204 MTIWEKEKLPPELOIREKNERYCYLVFNDEHHSYDHVYISLQALDCLAEALHTTAT 263

Db 1 MTIWEKEKLPPELOIREKNERYCYLVFNDEHHSYDHVYISLQALDCLAEALHTTAT 60

Qy 264 DKEGRAVKAAGAACQAEKEDIKSHSENVSHPLHVEVLHSEIMAHQFALRLGSMWNK 323

Db 61 DKEGRAVKAAGAACQAEKEDIKSHSENVSHPLHVEVLHSEIMAHQFALRLGSMWNK 120

Qy 324 IWSYSSDFQIFQACLRPEPSENPCLISRLMLWDAKLYKGARKILHILFSSPFMEWE 383

```

Db 121 IMSYSDPRQIFQOACLRBPDPGENPCLISRLMLWDAKLYKGARKILHELIFSSFFMEME 180
Qy 384 YKLFAMEFVKYKQLOKEYISDDHRSISITALSQOMFTVPTLARHLIEQNVISVITE 443
Db 181 YKLFAMEFVKYKQLOKEYISDDHRSISITALSQOMFTVPTLARHLIEQNVISVITE 240
Qy 444 TLEVLPEYLDNRNKNFQGYSDQKLRVYAVICDLKYILISKPTIWTIERLMOFLEGFR 503
Db 241 TLEVLPEYLDNRNKNFQGYSDQKLRVYAVICDLKYILISKPTIWTIERLMOFLEGFR 300
Qy 504 SFLKILTCMQMBEIRQVQCHIEVDPDWEAAIAIQMLXNILLMFQEWACDEBELLVA 563
Db 301 SFLKILTCMQMBEIRQVQCHIEVDPDWEAAIAIQMLXNILLMFQEWACDEBELLVA 360
Qy 564 YKECHKAVMRCSTSFISSTVTVQSCGHSLETYSYRVEDLVSIHPLSRTLAGLHVRLS 623
Db 361 YKECHKAVMRCSTSFISSTVTVQSCGHSLETYSYRVEDLVSIHPLSRTLAGLHVRLS 420
Qy 624 RLGAVERLHEFVDFQVEVLVEYPLRCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 683
Db 421 RLGAVERLHEFVDFQVEVLVEYPLRCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 480
Qy 684 REEMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTIISTKQDOLIKQNTLIEEM 743
Db 481 REEMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTIISTKQDOLIKQNTLIEEM 540
Qy 744 LQVLIYIVGERYVPGVGNVTKBEVWTREIILHLCIEPMPHSAIAKNLPENENNETGLENV 803
Db 541 LQVLIYIVGERYVPGVGNVTKBEVWTREIILHLCIEPMPHSAIAKNLPENENNETGLENV 600
Qy 804 INKVATFKPGVSGHGVYELKDESLKDFNMVYHYSTQHSKAEHMOKKRRKOENKDEAL 863
Db 601 INKVATFKPGVSGHGVYELKDESLKDFNMVYHYSTQHSKAEHMOKKRRKOENKDEAL 660
Qy 864 PPPPPPEFCFPAFSKVINLLNCDIMMYILRTVFERAIDTDSNLWTEGMLQWAPHILALGLL 923
Db 661 PPPPPPEFCFPAFSKVINLLNCDIMMYILRTVFERAIDTDSNLWTEGMLQWAPHILALGLL 720
Qy 924 EEKQOLOKAPEEVTFDFYHKASRLGSSAMNIOHLEKLGIPQLEGQKDMITWILQMPD 983
Db 721 EEKQOLOKAPEEVTFDFYHKASRLGSSAMNIOHLEKLGIPQLEGQKDMITWILQMPD 780
Qy 984 TVKRLREKSLIVATTSGSESINKNDEITHDK 1014
Db 781 TVKRLREKSLIVATTSGSESINKNDEITHDK 811

```

Search completed: July 9, 2005, 15:28:52  
 Job time : 141 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 14:29:13 ; Search time 49 Seconds  
(without alignments)  
2664.515 Million cell updates/sec

Title: US-10-758-672A-2  
Perfect score: 9224  
Sequence: 1 MADEAGGTERMEISAEPLQ.....BIARQETNQLFGFNWQLL 1749

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	4	US-09-724-126A-2
2	9087	98.5	1734	4	US-09-724-126A-19
3	8617	93.4	1757	4	US-09-724-126A-15
4	4327	46.9	1755	4	US-09-724-126A-4
5	4278	46.4	1755	4	US-09-724-126A-6
6	456	4.9	223	4	US-09-270-767-46327
7	450.5	4.9	334	4	US-09-270-767-45019
8	317.5	3.4	817	4	US-09-248-796A-19402
9	274.5	3.0	631	4	US-09-248-796A-19403
10	197	2.1	280	4	US-09-270-767-32482
11	190.5	2.1	3259	4	US-09-949-016-6507
12	184	2.0	1979	4	US-09-949-016-6468
13	184	2.0	2047	4	US-09-949-016-7404
14	176	1.9	2662	4	US-09-595-684B-31
15	175.5	1.9	1427	4	US-09-538-092-1044
16	175.5	1.9	2663	4	US-09-538-092-1252
17	172.5	1.9	1780	4	US-09-949-016-6899
18	172.5	1.9	1786	4	US-09-949-016-7880
19	172	1.9	1055	4	US-09-949-016-9776
20	172	1.9	3878	4	US-09-914-253-11
21	171	1.9	994	4	US-09-949-016-6779
22	164	1.8	2954	4	US-09-150-867-1
23	159.5	1.7	2733	4	US-09-949-016-11433
24	157.5	1.7	2789	4	US-09-949-016-8208
25	156.5	1.7	2710	1	US-08-480-604A-6
26	156.5	1.7	2710	2	US-08-405-496A-6
27	156.5	1.7	2710	3	US-08-915-136-6

28	156.5	1.7	2710	3	US-08-957-310-6	Sequence 6, Appli
29	156.5	1.7	2710	4	US-10-011-366-6	Sequence 6, Appli
30	156.5	1.7	2710	4	US-09-084-517-6	Sequence 6, Appli
31	155.5	1.7	2704	4	US-09-538-092-1260	Sequence 1260, Ap
32	152.5	1.7	3433	4	US-09-538-092-1136	Sequence 1136, Ap
33	150.5	1.6	1388	3	US-09-572-191-2	Sequence 2, Appli
34	150.5	1.6	1388	3	US-09-723-262-2	Sequence 2, Appli
35	150.5	1.6	1388	3	US-09-723-219-2	Sequence 2, Appli
36	149	1.6	3433	4	US-09-091-501B-10	Sequence 10, Appl
37	148	1.6	2125	4	US-09-919-172-29	Sequence 29, Appl
38	148	1.6	2871	4	US-09-538-092-936	Sequence 936, App
39	146.5	1.6	973	4	US-09-392-714-24	Sequence 24, Appl
40	146	1.6	958	4	US-09-171-991-6	Sequence 6, Appli
41	145.5	1.6	1401	4	US-09-750-590A-2	Sequence 2, Appli
42	144.5	1.6	1530	4	US-09-949-016-6668	Sequence 6668, Ap
43	142.5	1.5	1362	4	US-09-949-016-7033	Sequence 7033, Ap
44	142.5	1.5	1530	4	US-09-976-594-736	Sequence 736, App
45	142.5	1.5	3210	4	US-09-538-092-1154	Sequence 1154, Ap

ALIGNMENTS

RESULT 1  
US-09-724-126A-2  
; Sequence 2, Application US/09724126A  
; Patent No. 6706505  
; GENERAL INFORMATION:  
; APPLICANT: Han, Hui-Quan  
; APPLICANT: Kwak, Keith  
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family  
; FILE REFERENCE: 01017735966A  
; CURRENT APPLICATION NUMBER: US/09/724,126A  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,211  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 2  
; LENGTH: 1749  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-126A-2

Query Match		100.0%	Score 9224;	DB 4;	Length 1749;
Best Local Similarity		100.0%	Pred. No. 0;		
Matches 1749;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MADBEAGGTERMEISAEPLQTPPQRLASWWDQVDYTTAFLLHHLAQLVPEIYPAEMDPDL 60			
Db	1	MADBEAGGTERMEISAEPLQTPPQRLASWWDQVDYTTAFLLHHLAQLVPEIYPAEMDPDL 60			
Qy	61	KOBSVQMSIETPLEWYLFGBDDPDLCEKXHGSAFQCGRVFKSGETTYSRCDCAIDPT 120			
Db	61	KOBSVQMSIETPLEWYLFGBDDPDLCEKXHGSAFQCGRVFKSGETTYSRCDCAIDPT 120			
Qy	121	CVLCHDCFDQSVHKNHRYKMTSTGGGFCDDGTETAWKTGFCVNHHPGRAGTIKENSRC 180			
Db	121	CVLCHDCFDQSVHKNHRYKMTSTGGGFCDDGTETAWKTGFCVNHHPGRAGTIKENSRC 180			
Qy	181	FLNEEVIQVARKIPPSVIKYVEMTWEERKELPELQIREKNERYCYVLFNDEHHSYDH 240			
Db	181	FLNEEVIQVARKIPPSVIKYVEMTWEERKELPELQIREKNERYCYVLFNDEHHSYDH 240			
Qy	241	VYISLQALDCELAQAQLHTTAIDKEGRRVAKAGYACQAKEDIKSHSENVSHPLHV 300			
Db	241	VYISLQALDCELAQAQLHTTAIDKEGRRVAKAGYACQAKEDIKSHSENVSHPLHV 300			
Qy	301	EVLHSEIMAHOKFALRLGSSWNKIMSYSSDFRQIFCOACLRPEEPDSENPCLISRLMLWA 360			
Db	301	EVLHSEIMAHOKFALRLGSSWNKIMSYSSDFRQIFCOACLRPEEPDSENPCLISRLMLWA 360			
Qy	361	KLYGARKILHILFSSFFMEMEYKGLFAMBFVKKYKQLQKEYISDDHRSISITALSVDQ 420			



Db 361 KUYGARKILHBLIFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSVQ 420  
QY 421 MFTVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNPNFQGYSDQKLGRIYAVICDLK 480  
Db 421 MFTVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNPNFQGYSDQKLGRIYAVICDLK 480  
QY 481 YLISKPTIWTBRLMQFLEGFRSFLKILTCMOGHEEIRROVQGHIEVDPDWEAAIAIOM 540  
Db 481 YLISKPTIWTBRLMQFLEGFRSFLKILTCMOGHEEIRROVQGHIEVDPDWEAAIAIOM 540  
QY 541 OLKNILLMFOEWACDEBELLVAYKECHKAVNRCSPTSSSKTVVQSCGHSLETKSYRV 600  
Db 541 OLKNILLMFOEWACDEBELLVAYKECHKAVNRCSPTSSSKTVVQSCGHSLETKSYRV 600  
QY 601 SEDLVSIHLPLSRTLAGLHVRLSRLGAVSRHLEFVSFEDQVEVLVREYPLRCLVLVAQV 660  
Db 601 SEDLVSIHLPLSRTLAGLHVRLSRLGAVSRHLEFVSFEDQVEVLVREYPLRCLVLVAQV 660  
QY 661 AEMWRNGLSLISQVFFYQDVKCREMYDKOIMLQIGASLMDPNKFLILVLQRYELABA 720  
Db 661 AEMWRNGLSLISQVFFYQDVKCREMYDKOIMLQIGASLMDPNKFLILVLQRYELABA 720  
QY 721 FNKTIKTDODLKOYNTLIEBMLQVLIYVGERVPGVGNVTKBEVTNREIHLCLIEP 780  
Db 721 FNKTIKTDODLKOYNTLIEBMLQVLIYVGERVPGVGNVTKBEVTNREIHLCLIEP 780  
QY 781 MPHSAIAKNLPENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMFYHYSK 840  
Db 781 MPHSAIAKNLPENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMFYHYSK 840  
QY 841 TQHSKAEHMQKRRKOENDEALPPPPPPPCPAFSKVINLNCDIMMYILITVPERAID 900  
Db 841 TQHSKAEHMQKRRKOENDEALPPPPPPPCPAFSKVINLNCDIMMYILITVPERAID 900  
QY 901 TDSNLWTEGMLQWAFHILALGLEEKQLOKAPPEBEVTFDVKASRLGSSAMNLOMLE 960  
Db 901 TDSNLWTEGMLQWAFHILALGLEEKQLOKAPPEBEVTFDVKASRLGSSAMNLOMLE 960  
QY 961 KLKGIPLQSGQDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDEITHDKEAERK 1020  
Db 961 KLKGIPLQSGQDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDEITHDKEAERK 1020  
QY 1021 RKAEBARLHROKIMAGSALQKNFIETHKLMYDNTSEMPCKEDSIMEESTPAVDYSRI 1080  
Db 1021 RKAEBARLHROKIMAGSALQKNFIETHKLMYDNTSEMPCKEDSIMEESTPAVDYSRI 1080  
QY 1081 ALGPKRGPSVTEKEVLTCLCOBEOBKLENNAMVLSACVOKSTALTQHRGKPIELSGEA 1140  
Db 1081 ALGPKRGPSVTEKEVLTCLCOBEOBKLENNAMVLSACVOKSTALTQHRGKPIELSGEA 1140  
QY 1141 LDPLFMDPLAYGTYTSGCHVNHAVCWQKYFEAVQLSSQQRHVDLPDLESGEYLCPIC 1200  
Db 1141 LDPLFMDPLAYGTYTSGCHVNHAVCWQKYFEAVQLSSQQRHVDLPDLESGEYLCPIC 1200  
QY 1201 KSLCNTVPIPIIPIQOKINSENADALQLLTARWITQTVLARIISGNIRHAKENPIPIF 1260  
Db 1201 KSLCNTVPIPIIPIQOKINSENADALQLLTARWITQTVLARIISGNIRHAKENPIPIF 1260  
QY 1261 FNQMGDSTLEPHSILSFGVSESIKYSNSIKEMVILPATIIVIRIGLKVPPDERDPRVPM 1320  
Db 1261 FNQMGDSTLEPHSILSFGVSESIKYSNSIKEMVILPATIIVIRIGLKVPPDERDPRVPM 1320  
QY 1321 TWSTCAFTTIOAIENLLGDEGKPLFGALQNRQHNGLKALMQFAVAQRITCPQVLIQKHLVR 1380  
Db 1321 TWSTCAFTTIOAIENLLGDEGKPLFGALQNRQHNGLKALMQFAVAQRITCPQVLIQKHLVR 1380  
QY 1381 LLSVLPNLIKSEDTPCLLSIDLPHVLVGVAVLAPPSLYWDDPDVLDQSPSSVSSNHLHYLFH 1440  
Db 1381 LLSVLPNLIKSEDTPCLLSIDLPHVLVGVAVLAPPSLYWDDPDVLDQSPSSVSSNHLHYLFH 1440  
QY 1441 LITMAHMLQILLTVDVTGLPLAQVEDSEAHSSAFPAEISQYTSIGSDIPGWYLWVS 1500

Db 1441 LITMAHMLQILLTVDVTGLPLAQVEDSEAHSSAFPAEISQYTSIGSDIPGWYLWVS 1500  
QY 1501 LKNGITPYLRCAALFFHYLLGVTPPEELHTNSABGEYSALCSYLSLPTNLFLFQEWYDT 1560  
Db 1501 LKNGITPYLRCAALFFHYLLGVTPPEELHTNSABGEYSALCSYLSLPTNLFLFQEWYDT 1560  
QY 1561 VRPLLQRCWADPALLNCLKOKNTVVRYPKRNLSIELPDDYSCLINQASHFRCPRSADE 1620  
Db 1561 VRPLLQRCWADPALLNCLKOKNTVVRYPKRNLSIELPDDYSCLINQASHFRCPRSADE 1620  
QY 1621 RKHPVLCLFCGAILCSQNICCOEIVNGEVCACIFHALHCGAGVCIFLKIRECRVVLVEG 1680  
Db 1621 RKHPVLCLFCGAILCSQNICCOEIVNGEVCACIFHALHCGAGVCIFLKIRECRVVLVEG 1680  
QY 1681 KARGCAYPAPYLDYGETDGLKEGNPLHLSREYRKLHLVWQOHCIIIEETARSOETNOM 1740  
Db 1681 KARGCAYPAPYLDYGETDGLKEGNPLHLSREYRKLHLVWQOHCIIIEETARSOETNOM 1740  
QY 1741 LFGFNWOLL 1749  
Db 1741 LFGFNWOLL 1749  
RESULT 2  
US-09-724-126A-19  
; Sequence 19, Application US/09724126A  
; Patent No. 6706505  
; GENERAL INFORMATION:  
; APPLICANT: Han, Hui-Quan  
; APPLICANT: Kwak, Keith  
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family  
; FILE REFERENCE: 01017/35966A  
; CURRENT APPLICATION NUMBER: US/09/724,126A  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,211  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 1734  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-126A-19  
Query Match 98.5%; Score 9087; DB 4; Length 1734;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1728; Conservative 0; Mismatches 1; Indels 20; Gaps 1;  
QY 1 MADEAGGTERMEISAEPLQTPQRLASWWDQVDTTAFHLHLAQLVPEIYFAEMDPDLE 60  
Db 6 MADEAGGTERMEISAEPLQTPQRLASWWDQVDTTAFHLHLAQLVPEIYFAEMDPDLE 65  
QY 61 KOEESVQMSIPTPLEWYLFGEEDPDTCLBKXHSAGFQICGRVFKSGEYTSYCRDCAIDPT 120  
Db 66 KOEESVQMSIPTPLEWYLFGEEDPDTCLBKXHSAGFQICGRVFKSGEYTSYCRDCAIDPT 125  
QY 121 CVLCHMDCTODSVHKHRYKMHMTSTGGCDGDTBAKTTGTCVNVNHEPGRAGTICKENSRC 180  
Db 126 CVLCHMDCTODSVHKHRYKMHMTSTGGCDGDTBAKTTGTCVNVNHEPGRAGTICKENSRC 185  
QY 181 PLNEBVIQVARKIPPSVIKYVEMVTIWEBEKELPPELOIREKNERYCYVLFNDEHSHVDH 240  
Db 186 PLNEBVIQVARKIPPSVIKYVEMVTIWEBEKELPPELOIREKNERYCYVLFNDEHSHVDH 245  
QY 241 VIYSIQRALDCELAQAQLHTTAIDKEGRRVAKAGAYAAQCEAKEDIKSHSENVSOHPLHV 300  
Db 246 VIYSIQRALDCELAQAQLHTTAIDKEGRRVAKAGAYAAQCEAKEDIKSHSENVSOHPLHV 305  
QY 301 EYVLSHEIMAHQKFAIRLGSWMNKINSYSDRFOIFCOACLRPEEPSENPCILSRMLWDA 360  
Db 306 EYVLSHEIMAHQKFAIRLGSWMNKINSYSDRFOIFCOACLRPEEPSENPCILSRMLWDA 365  
QY 361 KLYKGARKILHBLIFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSVQ 420

Db 366 KLYGARKILHELIFSSPFMEYKCLFAMBFVKYKQKQKEYISDDHRSISITALSQ 425  
Qy 421 MFTVPTLARHLIEQNVSIVITETLLEVLPEYLDNRNKFNFQYSDQKLGRIYAVICDLK 480  
Db 426 MFTVPTLARHLIEQNVSIVITETLLEVLPEYLDNRNKFNFQYSDQKLGRIYAVICDLK 485  
Qy 481 YLISKPTIWTBRLMOPLEGFSPLKILTCQGEIEIRQVQHIETVDPDWEAAIAIQM 540  
Db 486 YLISKPTIWTBRLMOPLEGFSPLKILTCQGEIEIRQVQHIETVDPDWEAAIAIQM 545  
Qy 541 QLNKILLMFEQWACDEEILLVAYKECHKAVNRCSFTSSKTVVQSCGHSLETYSYRV 600  
Db 546 QLNKILLMFEQWACDEEILLVAYKECHKAVNRCSFTSSKTVVQSCGHSLETYSYRV 605  
Qy 601 SEDLSIHLPLSRTLAGLHVRLSRGAVSRLEHFEVSPFQVEVLVEYPLRCLVLVAQV 660  
Db 606 SEDLSIHLPLSRTLAGLHVRLSRGAVSRLEHFEVSPFQVEVLVEYPLRCLVLVAQV 665  
Qy 661 AEWWRNGLSLSQVFPYQDVKCRBEMDKDIIIMLQIGASLMDPNKFLLLVLRVELAE 720  
Db 666 AEWWRNGLSLSQVFPYQDVKCRBEMDKDIIIMLQIGASLMDPNKFLLLVLRVELAE 725  
Qy 721 FNKTISTKQODLIIKQNTLIEBMLQVLIYVIGERYVPGVGNVTKBEVTWREIHLCTIEP 780  
Db 726 FNKTISTKQODLIIKQNTLIEBMLQVLIYVIGERYVPGVGNVTKBEVTWREIHLCTIEP 785  
Qy 781 MPHSAIKNLPENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMYFYYSK 840  
Db 786 MPHSAIKNLPENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMYFYYSK 845  
Qy 841 TQSKAEHMKQRKQENKDEALPPPPPEPCFAPSKVINLNCDIMMYILTVFERAID 900  
Db 846 TQSKAEHMKQRKQENKDEALPPPPPEPCFAPSKVINLNCDIMMYILTVFERAID 905  
Qy 901 TDSNLTWTEGMLQWPHIALGALLESKQLOKAPBEVTFDFYHKASRLGSSAMNQLMLE 960  
Db 906 TDSNLTWTEGMLQWPHIALGALLESKQLOKAPBEVTFDFYHKASRLGSSAMNQLMLE 965  
Qy 961 KLKGPQLGQKDMITWILQMFDTVKRLREKSLIVATTSGSESINKDEITHDEKAERK 1020  
Db 966 KLKGPQLGQKDMITWILQMFDTVKRLREKSLIVATTSGSESINKDEITHDEKAERK 1025  
Qy 1021 RAEAAARLHRQIKMAQMSALQKNFTIETHKLYDNTSEMPGKEDSIMEBESTPAVSDYSRI 1080  
Db 1026 RAEAAARLHRQIKMAQMSALQKNFTIETHKLYDNTSEMPGKEDSIMEBESTPAVSDYSRI 1085  
Qy 1081 ALGPKRGPSTKEVLTCILCOEEOVKIENNAMVLSACVOKSTALTQHRGKPIELSGEA 1140  
Db 1086 ALGPKRGPSTKEVLTCILCOEEOVKIENNAMVLSACVOKSTALTQHRGKPIELSGEA 1145  
Qy 1141 LDPLFMDPLAYGTYTSGCGHVMHACVQKYPEAVQLSSQQRHVDLPDLESGLYCLPLC 1200  
Db 1146 LDPLFMDPLAYGTYTSGCGHVMHACVQKYPEAVQLSSQQRHVDLPDLESGLYCLPLC 1205  
Qy 1201 KSLCNTVPIIPLQPKINSENADALQALLTLARWITQVLARISGYNIRHAKGENPIPIF 1260  
Db 1206 KSLCNTVPIIPLQPKINSENADALQALLTLARWITQVLARISGYNIRHAKGENPIPIF 1265  
Qy 1261 FNQMGDSTLEPHSILSRGVSSIKYSIKEMVILPATIIVRIGIKVPPDPRVPM 1320  
Db 1266 FNQMGDSTLEPHSILSRGVSSIKYSIKEMVILPATIIVRIGIKVPPDPRVPM 1325  
Qy 1321 TWSTCAFTIQTALNLLGDEGKLPFALQNRQNGKALKMQPVAQRITCPOVLIOKHLVR 1380  
Db 1326 TWSTCAFTIQTALNLLGDEGKLPFALQNRQNGKALKMQPVAQRITCPOVLIOKHLVR 1385  
Qy 1381 LLSVVLPNIKSBDTFCLLSIDLPHVLVGAFLAPPSLYMDDPVDLPQSSVSSSYNHLFLH 1440  
Db 1386 LLSVVLPNIKSBDTFCLLSIDLPHVLVGAFLAPPSLYMDDPVDLPQSSVSSSYNHLFLH 1445  
Qy 1441 LITMAHMLQILLTVDTGLPLAQVQDSEBHSASSFPABISQYTSIGSCDIPGHWLWS 1500

Db 1446 LITMAHMLQILLTVDTGLPLAQVQDSEBHSASSFPABISQYTSIGSCDIPGHWLWS 1505  
Qy 1501 LKNGITPYLRCAALPFHYLLGVTTPBELHTNSABEYSALCSYLSLPTNLFLFLFOBYMDT 1560  
Db 1506 LKNGITPYLRCAALPFHYLLGVTTPBELHTNSABEYSALCSYLSLPTNLFLFLFOBYMDT 1565  
Qy 1561 VRPLLQRCADPALANCLKOKNTVYRPRKNSLIELPDDYSCLLNOASHFRCPRSADDE 1620  
Db 1566 VRPLLQRCADPALANCLKOKNTVYRPRKNSLIELPDDYSCLLNOASHFRCPRSADDE 1625  
Qy 1621 RHGPVLCIFCGAILCSQNICCOEIVNGEVEGACIFALHCGAGVCIFLKIREFCRVVLVEG 1680  
Db 1626 RHGPVLCIFCGAILCSQNICCOEIVNGEVEGACIFALHCG 1685  
Qy 1681 KARGCAYPAPYLDREYGETDPGLKGNPLHLSRERYRKLHLVWQHCIIIEIARSQETNOM 1740  
Db 1686 KARGCAYPAPYLDREYGETDPGLKGNPLHLSRERYRKLHLVWQHCIIIEIARSQETNOM 1725  
Qy 1741 LFGFNWQLL 1749  
Db 1726 LFGFNWQLL 1734

## RESULT 3

US-09-724-126A-15

; Sequence 15, Application US/09724126A

; Patent No. 6706505

; GENERAL INFORMATION:

; APPLICANT: Han, Hui-Quan

; APPLICANT: Kwak, Keich

; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family

; FILE REFERENCE: 0107735966A

; CURRENT APPLICATION NUMBER: US/09/724,126A

; CURRENT FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US 60/187,211

; PRIOR FILING DATE: 1999-03-01

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patent version 3.0

; SEQ ID NO 15

; LENGTH: 1757

; TYPE: PRT

; ORGANISM: Mouse

US-09-724-126A-15

Query Match 93.4%; Score 8617; DB 4; Length 1757;

Best Local Similarity 92.1%; Pred. No. 0;

Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;

Qy 1 MADEBAGTERMEBISAEPLPQTPORLASWWDQOVDYFATFLHHLAQVLPEIYFAEMDPDLE 60  
Db 1 MADEBMDGAERMDVSPPEPLAPQRPASWWDQOVDYFATFLHHLAQVLPEIYFAEMDPDLE 60  
Qy 61 KOESVQMSIITPLEWYLFGRDPDLCLEKLXHGAFQLCGRVFKSGETTSYSCRDCAIDPT 120  
Db 61 KOESVQMSIITPLEWYLFGRDPDLCLEKLXHGAFQLCGRVFKSGETTSYSCRDCAIDPT 120  
Qy 121 CVLQMDCFQSDVHKHNRHYMTSTGGGFCDCGTEAWKTGPPFCVNHPEGRAGTITKENSRC 180  
Db 121 CVLQMDCFQSDVHKHNRHYMTSTGGGFCDCGTEAWKTGPPFCVNHPEGRAGTITKENSRC 180  
Qy 181 PLNBEVIVQARKIPSVUTKYVEMTWEERKELEPPELOIREKNERYCYVLFNDEHSHVDH 240  
Db 181 PLNBEVIAQARRIPFSVIKYIVEMTWEERKELEPPELOIREKNERYCYVLFNDEHSHVDH 240  
Qy 241 VIYSIQRALDCELAQAOLHTTAIDKEGPRVAKAGYACQAKEDIKSHSENSVSHPLHV 300  
Db 241 VIYSIQRALDCELAQAOLHTTAIDKEGPRVAKAGYACQAKEDIKSHSENSVSHPLHV 300  
Qy 301 EVLHSEIMAHQKFAIRLGSMWNKIMSYSSDPQFCQACLEEDPSENPCILSRMLMWA 360  
Db 301 EVLHSEIMAHQKFAIRLGSMWNKIMSYSSDPQFCQACLEEDPSENPCILSRMLMWA 360  
Qy 361 KLYGARKILHELIFSSPFMEYKCLFAMBFVKYKQKQKEYISDDHRSISITALSQ 420

Db 361 KLYKGARKLHLLHIFSSFWMEYKGLFAMEFVKYKQLQKEYISDDHRSISITALSQ 420  
Qy 421 MFTVPTLARBHLIEEQNVISVITETLLEVLPEYLDNNKFNPOGYSQDKLGRVYAVICDLK 480  
Db 421 MFTVPTLARBHLIEEQNVISVITETLLEVLPEYLDNNKFNPOGYSQDKLGRVYAVICDLK 480  
Qy 481 YLISKPTTWTERLRMOLEGRSPFLKILTCMOGHEETIRROVGOHLEVDPDWEAAIAIQM 540  
Db 481 YLISKPTTWTERLRMOLEGRSPFLKILTCMOGHEETIRROVGOHLEVDPDWEAAIAIQM 540  
Qy 541 QLNKILLMFOEWACADELLVAYKECHKAVNRCSFTSSKTVVQSGHSLKTSYRV 600  
Db 541 QLNKILLMFOEWACADELLVAYKECHKAVNRCSFTSSKTVVQSGHSLKTSYRV 600  
Qy 601 SEDLSIHLPLSRTLAGLHVRLSRGAVSRHLEFVSFQVEVLVEYPLRCLVIAQV 660  
Db 601 SEDLSIHLPLSRTLAGLHVRLSRGAVSRHLEFVSFQVEVLVEYPLRCLVIAQV 660  
Qy 661 AEMWRNGLSLISQVYQDVKREMYDKDIIMLOIGASLMDPNKFLILVLRVELAEA 720  
Db 661 AEMWRNGLSLISQVYQDVKREMYDKDIIMLOIGASLMDPNKFLILVLRVELAEA 720  
Qy 721 ENKISTKQDILIKONTLIEEMQLVLIYIGERYVPGVGNVTKBEVTREIHLCLIEP 780  
Db 721 ENKISTKQDILIKONTLIEEMQLVLIYIGERYVPGVGNVTKBEVTREIHLCLIEP 780  
Qy 781 MPHSAIARLNENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMYFYYSK 840  
Db 781 MPHSAIARLNENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMYFYYSK 840  
Qy 841 TQHSKAEHMKRRKQENKORALPPPPPPPCPAPSKVINLANCDIMWYILVTPERAID 900  
Db 841 TQHSKAEHMKRRKQENKORALPPPPPPPCPAPSKVINLANCDIMWYILVTPERAID 900  
Qy 901 TDSNLTWTEGLOWAPHILALGILLERKQOLKAPPEERVTDFYHKASRLGSSAM--NIQM 957  
Db 901 TDSNLTWTEGLOWAPHILALGILLERKQOLKAPPEERVTDFYHKASRLGSSAM--NIQM 957  
Qy 958 LLEKLGIPQLEGQKDMITWILQMPDVTVRLEKSLIVATTSGSISIKNDITHDKKA 1017  
Db 958 LLEKLGIPQLEGQKDMITWILQMPDVTVRLEKSLIVATTSGSISIKNDITHDKKA 1017  
Qy 1018 ERKRKAABARLRQKIMQMSALQKQFIETHKLMDYNTSEMPGKEDSINEERESTAVSDY 1077  
Db 1018 ERKRKAABARLRQKIMQMSALQKQFIETHKLMDYNTSEMPGKEDSINEERESTAVSDY 1077  
Qy 1078 SRIALGPKRGPSVTEKEVLTCILCQEEQVKIENNAWLSACVQKSTALTQHRGKPIELS 1137  
Db 1078 SRIALGPKRGPSVTEKEVLTCILCQEEQVKIENNAWLSACVQKSTALTQHRGKPIELS 1137  
Qy 1138 GERALDPLFMDPDLAYGTYTSGCHVNHAVCWQKYPEAVQLSSQRIHVDLPDLESGEYLC 1197  
Db 1138 GERALDPLFMDPDLAYGTYTSGCHVNHAVCWQKYPEAVQLSSQRIHVDLPDLESGEYLC 1197  
Qy 1198 PLCKSLCNTVPIPILOPKINSENADALQILLTLARWITQTLARI SGYNIRHAKENP- 1256  
Db 1198 PLCKSLCNTVPIPILOPKINSENADALQILLTLARWITQTLARI SGYNIRHAKENP- 1256  
Qy 1201 PLCKSLCNTVPIPILOPKINSENAEALQILLTLARWITQTLARI SGYNIRHAKGEAPA 1260  
Db 1201 PLCKSLCNTVPIPILOPKINSENAEALQILLTLARWITQTLARI SGYNIRHAKGEAPA 1260  
Qy 1257 IPIFFNQMGDSFLRPHSILSPGVSSIKYSNISIKEMVILFATTIRIGLKVPPDPR 1316  
Db 1257 IPIFFNQMGDSFLRPHSILSPGVSSIKYSNISIKEMVILFATTIRIGLKVPPDPR 1316  
Qy 1317 VPMLTWSTCAFTIQAENILGDBGKPLFGALQNRQHNGKALMQFAVQIRITCPQVLIQK 1376  
Db 1317 VPMLTWSTCAFTIQAENILGDBGKPLFGALQNRQHNGKALMQFAVQIRITCPQVLIQK 1376  
Qy 1377 HLVRLLSVLPNKTSEDTPCLLSIDLPHVLVGNVLA PPSLYWDDPDVLOPSSVSSYNHL 1436  
Db 1377 HLVRLLSVLPNKTSEDTPCLLSIDLPHVLVGNVLA PPSLYWDDPDVLOPSSVSSYNHL 1436  
Qy 1437 YLFHLITMAHMLQILLTVDVTDV---GLPLAQVQSDSEAHSSAFBISQVTSIGGCDI 1492  
Db 1437 YLFHLITMAHMLQILLTVDVTDV---GLPLAQVQSDSEAHSSAFBISQVTSIGGCDI 1492

Db 1441 YLFHLITMAHMLQILLTDTDLSPGPPLAEGEEDSEARCASAFFVEVSOHTDGLTGCGA 1500  
Qy 1493 PGWYLVSLKNGITPYLRCAALFPHYLLGVTPPPELHTNSAEGEVSALCSYLSLPTNLPL 1552  
Db 1501 PGWYLVSLKNGITPYLRCAALFPHYLLGVTPPPELHTNSAEGEVSALCSYLSLPTNLPL 1560  
Qy 1553 LFOEYWDVTRPDLQWACADPALLNCLKQKNTVVRYPKRNSLIELPDDYSCLLNQASHFR 1612  
Db 1561 LFOEYWDVTRPDLQWACADPALLNCLKQKNTVVRYPKRNSLIELPDDYSCLLNQASHFR 1620  
Qy 1613 CPRSADDERKHPVLCLFCGAILCSQNICQBI NVGEVACIFHALHCGAGVCIPLKIRE 1672  
Db 1621 CPRSADDERKHPVLCLFCGAILCSQNICQBI NVGEVACIFHALHCGAGVCIPLKIRE 1680  
Qy 1673 CRVVLVEKARGCAYPAPYLDYGETDPLKRGNPLHLRSERYRKLHLVMOQHCHIEITA 1732  
Db 1681 CRVVLVEKARGCAYPAPYLDYGETDPLKRGNPLHLRSERYRKLHLVMOQHCHIEITA 1740  
Qy 1733 RSQETNOMLPGFNWQLL 1749  
Db 1741 RSQETNOMLPGFNWQLL 1757

## RESULT 4

US-09-724-126A-4

; Sequence 4, Application US/09724126A

; Patent No. 6706505

; GENERAL INFORMATION:

; APPLICANT: Han, Hui-Quan

; APPLICANT: Kwak, Keith

; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family

; FILE REFERENCE: 01017/35966A

; CURRENT APPLICATION NUMBER: US/09/724,126A

; CURRENT FILING DATE: 2000-11-28

; PRIOR FILING DATE: 1999-03-01

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 1755

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-126A-4

Query Match 46.9%; Score 4327; DB 4; Length 1755;

Best Local Similarity 47.2%; Pred. No. 0;

Matches 841; Conservative 339; Mismatches 543; Indels 58; Gaps 27;

Qy 1 MADE---EAGGTTER--MEISAE LPTQRPQRLASWMDQVDFTYTAFLHLAQLVPEIYFAEM 55  
Db 1 MASELEPEVQAIDRSLLLECSAE-----EIAQKWLQATDLTEVYQHLAHPVPIYCRGP 54  
Qy 56 DPDLKEQESVQMSIFTEPLEWYLFCEDPDI CLEKLKHSG-AFQLCGRVFKSGETTYSCRD 114  
Db 55 NPFPQKEDMLAQHVLGLPMWYLCGEDPAFGPKLEQANKPSHLCGRVFKVGEPTYSCRD 114  
Qy 115 CAIDPTCVLCHDPCFQDSVHKHRYKMYHTSTGGFCDCGDTAWKTGPPCVNHEPGRAG-T 173  
Db 115 CAVDPTCVLCECFPGSITHRDHYRMTTSGGGFCDCGDTAWKGGPYCQKHELTSEIE 174  
Qy 174 IKENSRCPLNEEVIQVARKIPSVIKYVYVMTIWEKEKELPELQIREKNERYYCVLPND 233  
Db 175 BEEDPLVHLSQEDVIARTNIFAIFRYAVAVILTWKESELPALEWKESTDYICMLFND 234  
Qy 234 EHSYDHYVYSIQRALDCELAQAHLHTTAIDKEGBRRAVAGAYAAQCBAKEDIKHSENV 293  
Db 235 EVHTYEQVIYTLQKAVNCTQKEAIGFAITVDRDGRSRYGDFQCEQAKSVIVRNTSRQ 294  
Qy 294 SOHPLHVEVLSEIWAHQKFAIRLGLSWMNKIMSYSDDSRQIFCQACLREEDSENPCILS 353  
Db 295 TK-PLKVQVMTSSIVAHQNFGLKLSLWLSIIGSDGLRILLCVQLGPGDGENSLVD 353  
Qy 354 RLMLWDAXLYKGARKILHELIFSSPFMEMEYKCLFAMEFVKYKQLQKEYISDDHRSIS 413

354 RLMLSDSKLWKGARSYTHQLFWSLLMDLKYKLPFAVRFAKNYQQJQDFMEDDHERAVS 413  
414 ITALSVMFTPTVPLARHLEEQNVISVITETLEVLPEYLDRNNKNFQOYS---QDKLG 470  
414 VTALSVOFTPTAPTLARMLTEENLMSIIKTMDHL-RHRDAQGRFOFERYTALQAFKR 472  
471 RVYAVICDLKYILI SKPTTWTERLMQPLEGFRSFLKILTCMQGMEIEIRVOQHLEVPD 530  
473 RVQSLILDLYVILSKPTWSDSLRQKFLGFPDAPLELLKCMQMDPITRQVQGHTEMEP 532  
531 DNEAATAIQMOLKNIILLMFOEWACDBELLVAYKECHKAVMRCSTSPITSSSKTVVQS-C 589  
533 EWEAAATLQWLTHVLSMWDWACASDEKVLIEAYKCLAVLMQCHGYTDGEPITLSIC 592  
590 GHSLETKSVRSBDLVSIHLPLSRTLAGLHVRLSRLGAVSRLEHFPSPDFQVEVLVREYP 649  
593 GHSVETIRYCVSQEKVSIHLPSVRLLAGLHVLSKSEVAYKPPPELLPLSELSPMLIEHP 652  
650 LRCLVLVAQVVAEMTRNGLSISQVFTYQDVKCREMYDKDIIMLOIGASLMDPKFL 709  
653 LRCLVLCAQVHAGMWRNGFSLVNIQIYYHNKCREMFQDVVLMQTVGSMMDPNHFLM 712  
710 LVLRVELAEPN-----KTISTK---DODLTKOYNLTLEEMQLVLIYIGERVVPGVNV 762  
713 IMLSRFELQIFSTPDYGRFSSEITHKDVQVQNNLTLEEMLYLIIMLVGERFSPGVQV 772  
763 TKEEVTMBRIIHLLEPIEMPHSAIAKNLPENENNETGLENNVINKVATFKKPGVSGHYVE 822  
773 NATDEIKREIIHQLSIKPMASSELVKSLEPENENNETGMESSIEVAHFKKPGLTGRGMYE 832  
823 LKDESILKDFNMFTYHYSKQHSKAEHMKRKRQKQENKDALPPPPPEPCAPSKVINL 882  
833 LKPECAKBFNLYFTHSRQAQKAEBAQRKLRQNRREDTALPPVLPFPCLPASLVNLT 892  
883 NCDIMMYILRTFERAIDTDSNLTGMLQWAPHILALGLLEKQOQKAPREE-VTPDP 941  
893 QSDVMLCITGTLQWAVEHNGYAWSESMLQVRLHLIGMALQOEKQHLNVTBEHVVTPT 952  
942 YHKASRLGSAWN---IOMLEKLGIPOLGOKOMITWILQWFDTVVGRLEKS-CLIVA 997  
953 TQKISKPGEPKNSPISLAWLETQNAPLYEVHKNIRWILKTFNAVKORSSPISPA 1012  
998 TTSGESIKNDIETHDKKAEARKKABARLHRQKIMQMSALQKNFIETHKLMYNTSE 1057  
1013 ETEGTTM---EESSRDKDAERKRKABARLHRQKIMQMSALQKNFIETHKLMYNTSE 1057  
1058 MPKEDSIMEESTRAVSDYSRLNLPKGPSPVTEKEVLTCTLCQEEQSVKLENNAMVLS 1117  
1070 LDASTSAVLDH--SPVASDMTLTALGPTQTPQEQRFVTCILCOEEQSVKLENNAMVLS 1127  
1118 ACQVKSTALTOHKGKPIELSGBALDPLFMDPLAYGTYTGSCHVNVHVCWKYFEAVOL 1177  
1128 AFQVQSTVLKRSKRFQI-DPEKYDPLFMDPLSCGTHTSCHGIMHACWQRYFDSVQA 1186  
1178 SQQ-----RIHVDLPDESGEYLCPLCKSLCNTVPIIPLQPKINSENADALAQLLPL 1232  
1187 KEQRRQRLRLHTS-YDVENGFEFLCPLCELSNTVPIPLL-LPPNFIENNRLN-FSDOPNL 1243  
1233 ARWIOVTLARISGYNIRHAKGENPDIFFNOQMGDSTLEPHILSPGVESSIKYNSIKE 1292  
1244 TQWIRTISOQIKALQFLRKEESTPNNASTKNSENVDELQPLGFRPDPRPKPIPYSESIKE 1303  
1293 WYLPATTYIRIGLVKPPDERPRVPMLTWSTCAFTIOAIENLGLDGEKPLFGALQNRQH 1352  
1304 MUTTGATYKVLGVHNEEDPRVIMCWSGCAVYIOIERILSDSDKPLGFLPCLRD 1363  
1353 NGLKALMQFAVQRICTCPQVLIQKHLVRLLSVVLPNIKSEDTPLCLSIDLFLVGLVLA 1412  
1364 DCLRSITRFAAAHTVAVSVVQGHFCLPASIVPNDSSHEELPCILIDMFHLLVGLVLA 1423  
1413 FPLSLYWDPPVDIOPSSVSSVSNHLYLFLHITMAHMLQILL-----TVDTLPLAQVQDSEE 1469

1424 FPALQCOD-----PSGISLGTGDLHLFHLVTVMAHIIQILLTSTCTENGMD--QENPPCBE 1476  
1470 AHSASFFABISQYTSIGSDIP-GWYLMVSLKNGITPYLRCAALFFHYLLGVTPPEL 1528  
1477 EBAVALALYKTLQYTG-SALKKEIPSGHLMWSVRAGIMPELFCALFFHYLLGVTPPEL 1535  
1529 HTNSAEGEYSALCSYLSLPTNLPLFLQBYWDTVRPLQBWCADPALLNCLGKQNTVVRP 1588  
1536 QV-PGTSHPHEHLCYLSLPTNLPLFLQBYWDTVRPLQBWCADPALLNCLGKQNTVVRP 1594  
1589 RKRNSLIELDDYSILLNOASHFRCPRSADDERKHIPVLCFLCGAILCSQNICQEIYNGE 1648  
1595 RESNKLINLPEDYSSILNOASHFRCPKSGDKSRAPTLCLVCGSLCSQSYCCQTELEGE 1654  
1649 EYVACIFHAIHCGAGVCIPLKIRCBVWLVEGKARGCAYPAPYLDYBGTDPGLKRGNPL 1708  
1655 DVGACHTAHTYSCSGVIGFLRVEBQVFLAGTKGCFSPYLDYBGTDPGLKRGNPL 1714  
1709 HLSRERYKHLVWQOHCIIIEIARSQETNOMLFGFNWQLL 1749  
1715 HLCKERFKKIQKWHQHSVTEIGHAQEANOVLGIDWQHL 1755

RESULT 5  
US-09-724-126A-6  
; Sequence 6, Application US/09724126A  
; Patent No. 6706505  
; GENERAL INFORMATION:  
; APPLICANT: Han, Hui-Quan  
; APPLICANT: Kwak, Keith  
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family  
; FILE REFERENCE: 0101735966A  
; CURRENT APPLICATION NUMBER: US/09724,126A  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,211  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 6  
; LENGTH: 1755  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-724-126A-6

Query Match 46.4%; Score 4278; DB 4; Length 1755;  
Best Local Similarity 46.5%; Pred. No. 0;  
Matches 835; Conservative 341; Mismatches 534; Indels 84; Gaps 28;

QY 1 MADE---BAGGTER--MISIAELPOTPORLASWQDQVDTAFHLHLAQLVPEIYFAEM 55  
DB 1 MASEMEPEVQAIDRSLLCSAE-----EIAGRWLQATDLNREVYQHLAHCVPKIYCRGP 54  
QY 56 DPDLKEQESVQMSIPTPLEWYLFCEDPDICLEKLKHSQ-AFOLCGRVPKSESTYSCHD 114  
DB 55 NPFPOKEDTLQAHLILGPMWYICABDPALGPKLEQANKPSHLGCRVPKVEPTYSCHD 114  
QY 115 CAIDPTCVLCMDCFQDSVKNHRYKMTSTGGFCDCGDTAWKTGPPCVNHEPGRAGTI 174  
DB 115 CAVDPTCVLCMECFQDSVKNHRYKMTSTGGFCDCGDTAWKTGPPCVNHEPGRAGTI 174  
QY 175 -KENSRCPLNBEVIVQARKIPSVIKYVYVMTIWEKEKELPELQIRKQERYVYCVLND 233  
DB 175 BEEDPLVHLSDEVIARTYNIIPALMRYAVDILTWKESSELPEDLSEVASKSDTYTCWLFND 234  
QY 234 BHSYDVHTYSLQALDCELAELQHTTAIDKEGRVAKAGAYACQAKEDIKSHENV 293  
DB 235 EVHTYEVIYTLQKAVNCTQKBAIGFATTVDRDRPRVRYGDFCYDQAKTVIVRNTSRQ 294  
QY 294 SOHLRHLVLEHSEIMAHQKVALRLGSMNKMYSVSSDFROIQACLRPEEPSENPLIS 353  
DB 295 TK-PLKVQVMSVAAHQVGLKALSLVGSVIGTSDGLRRLILQVGLQEGPGEENSLVD 353  
QY 354 RLMLWDAKLYKGARKILHELIFSSPFMEYKCLFAMBFVKYKQKQKEYISDDHRSIS 413

Db 354 RLMNDKLUWKGARSVYHQLFMSLLMDUKYKKLPAIRPAKRYRQLQDFWEDDHERAVS 413  
Qy 414 ITALSVMQFTVPTLARHLIEQNVSIVITETLLEVLPEYLDNRNKNFPGYS---QDKLG 470  
Db 414 VTALSVMQFTVPTLARMLLTENLMTVIIKAFMDHL-KHRDAQGRFQPERYATALQAKFR 472  
Qy 471 RYAVATCDLYILISKPTTWTBRLRMQFLEGFRSLKILTCHQGBEIRBQGHIEVDP 530  
Db 473 RYQSLTLDLYILISKPTTWSDELQKFLQGFDAFLLELCKMCGMDPTTRQVQGHIEBP 532  
Qy 531 DREAALAIQOMLKNILLMQEACDELLLVAYKECHKAVNRCSFSSSKTVQVS-C 589  
Db 533 EWEAAFTLMQLTHVISMVQDMCALDERKILIBAYKKCLAVLTQCHGFTDGEQPTLSIC 592  
Qy 590 GHSLETKSRVSEDLVSIHLPLSRTLAGLHVLRLSRLGAVSRHLHFVSEDFQVEVLVEYP 649  
Db 593 GHSVETIRVCVQEKVSIHLPIISRLLAGLHVLRLSSEVAYKPELLPLSELSPMMLIEHP 652  
Qy 650 LRCLVLVAQVAMRRNGLSLSQVFTYQDVKCBEMVDKDIIMLOIGASLMDPKELL 709  
Db 653 LRCLVLCAQVHAGMWRNGFSLVNOIYYHYHVKCREMPDKDIIVMLQTVSMMDPNHFLM 712  
Qy 710 LVLQRYELAEAFN----KTIETK--DODLIKOYNTLIBEMLQVLIYIVGERYVPGVNV 762  
Db 713 IMLSRPELQLFSTPDYGRKRFSEVTHKOVVQNNLIBEMLYLIIMLVGERFNGVGV 772  
Qy 763 TKEEVTMRIIHLICIEPMPSHAIKALNPENNETGLENVINKVATFKKPGVSGHGYE 822  
Db 773 AATDEIKREIIHOLSIKPAHSELVKSLEPENKETGMESVIESVAHPKPKGLTGRGMYE 832  
Qy 823 LKDESILKDNMTFYHVKTHSKAEHMQKRRKQENKDEALPPPPPEPCPAFSKVINLL 882  
Db 833 LKPECAKEPNLYFYHFSRAEQSKAEBAQRKLRKENKEDTALPPPALPPCPLFASLVNLL 892  
Qy 883 NCDIMMYILRTVFERAIDTDSNLWTEGMLQMAFHILALGLLEKQKQLOKAPREEV-TPDF 941  
Db 893 QCDVMUYIMGTILQWAVEHGHGSAWSEMLQVRLHLIGMALQEEKHLENAVGHVQVTF 952  
Qy 942 YHKASRLGSSAMN---IQMLLEKLGIPQLEGOKMITWILQMDPTVVKRLREKSLIVAT 998  
Db 953 TQKISKPGADPHNSPSILAMLETQNAPSLEAHKOMIRMLLMFNAIKKIRE--CSSSP 1010  
Qy 999 TGSSEIKNDEITHDKEKARKKABAAIRHOKMAQMSALQKNFIETHKLMYNTSM 1058  
Db 1011 VAAEAGTMEESSRDKDKAERKKAIRLRREKIMQAQSEMRHFIDENKELFOQTLEL 1070  
Qy 1059 PGKEDSIMEESTPAVDYSRTALGPKRGSVTEKEVLTCILQEBQYVKIENNAWLSA 1118  
Db 1071 DTSASATL--DSSPPVSDAALTALGPAQTVQPEPROFVTCILQEBQVTVGSRAMVLA 1128  
Qy 1119 CVQKSTALQHRKPIELSGEALDPLFMDPLAYGTYTTCGSHVMHVAWCQKPYEAVQUS 1178  
Db 1129 FVQBSTVLKORTKTI-ADPEKYDPLFMHPLDSCGTHTCGSHVMHACHWQRYFDSVQAK 1187  
Qy 1179 SQQ-----RIHVDLPLESGEYLCPLCKSLCNVTPIIPLOKQKINSENADALQLLTIA 1233  
Db 1188 EQRROQRLRHTS-YDVENGFEPLCECLSNVTPIPL-LPPRSILSRRLN-FSDQPDJA 1244  
Qy 1234 RWTQTLARISGNI---RHANGEN-----PIPIPNQMGDSTLEFHSILSF 1278  
Db 1245 QMTRATVQKVVQVURRKHNAADTSSSEDTAMNIIPIPEGFRP-----DFY----- 1292  
Qy 1279 GVSESIKYSNIKEMVILFATTIYRIGLKVPPDERPRVPMLTWTMSTCAFTIQAENLGD 1338  
Db 1293 ---PRNPYSDSIKEMLTITGCTAAKYGLKVHPNEDGPRVPILCWGTCAVTIQSIERILD 1349  
Qy 1339 EGKPLFGALQNRHNGLKALMQPAVAQRITCPOVLQKHLVRLSVVLNPKISEDTPCLL 1398  
Db 1350 EEPFVGPFLPCRLDDCLRSITREPAAAHWTVALLPVVQGHFKCLFASLVPSDSYEDLPCLT 1409  
Qy 1399 SIDLPHVLGAVLAFPSLYWDDPVDLQPSVSSVSNHLYLFLITWAHMLQILL---TVD 1455

Db 1410 DIDPHLLVLVLAPPALQCOD-----PSGSSLATGDLHIHFLVTAHIVQILLTCTEE 1464  
Qy 1456 TGLPLAQVQESBEAHSASSFABISQYTSIGCDIPGWTYLVWSLKNGIPTPYLCAALF 1515  
Db 1465 NGMD--QENPTGEBELAILSLHKTILHQYTGSALEAPSGMHLWRSVRAAIMPFLLKCSALF 1522  
Qy 1516 PHYLGVTPPELHTNSAEGYSALCSYLSLPTNLFLLPQBYWDTVRPLLRQWCADPALL 1575  
Db 1523 PHYLNGVPAPDQLV-SGTSHEFLCNLYLSLPTNLHILFQENSIDMNSLIESWCQNSEVK 1581  
Qy 1576 NCLQKQNTVVRYPKRNLSIELPDDYSCLLNAQASHFRCPRSADDERKHPVLCLFCGAILC 1635  
Db 1582 RYLNGERGAISYPRGANKLIDLPEYSSLIQAQSNFSCPKSGGDKSRAPTLCLVCGSLIC 1641  
Qy 1636 SONTCCQBIIVGEEVGACIFHALHCGAGVCFILKIRECRVVLVEGKARGCAYPAPYLB 1695  
Db 1642 SOSYCCQAELEGEDVGACTAHTYSCGAGIFLRVRECVLFLAGTKGCFYPPYLDY 1701  
Qy 1696 GETDPLKRGNPLHLRERYRKLHLVMOQHCIEBIARSQETNOMLFGFNWOLL 1749  
Db 1702 GETDGLRGNPLHLQCERFRKIQKLMQOHSITEIGHAQEANKOTLVGIDWQHL 1755

## RESULT 6

US-09-270-767-46327  
; Sequence 46327, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 46327

; LENGTH: 223

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-09-270-767-46327

Query Match 4.9%; Score 456; DB 4; Length 223;

Best Local Similarity 40.6%; Pred. No. 1.9e-32;

Matches 95; Conservative 40; Mismatches 85; Indels 14; Gaps 4;

Qy 750 IVERYPVGVGVNTKVEVTMBEIIHLLCIEBPMPSHAIKALNPENB--NNETGLENVINKV 807

Db 1 IIGERWMPGVSVTDEDRLEKKEIIQLLCIKPYSHSELRALPDGNSGSDNVFEEVINTV 60

Qy 808 ATPKKP-GVSGHGVYELADESLKDFNMYFYHYSKTHSKAEHMQKRRKQENKDEALPPP 866

Db 61 AVFKKPVGADSKGVYELKEHLKKEFNMYFYHYTKEDSKABELQERRAKKQLVCCPPP 120

Qy 867 PPEPCPAFSKVINLLNCDIMMYILRTVFERAIDTDSNLWTEGMLQMAFHILALGLLEK 926

Db 121 MLPLKLTAPTMANILQCPVPLNCLSLMERALNAYSRSPTESHQKVLHLGLYAIQES- 179

Qy 927 QOLQKAPREEVTFDPYHKASRLGSSAMNIQMLLEKLGIPQLEGOKMITWILQ 980

Db 180 -----LSEHYPLSFYSRSQBYG-----ILEKLRELCRPLEAHYDFVLTWIE 223

## RESULT 7

US-09-270-767-45019

; Sequence 45019, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517



SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 45019

LENGTH: 334

TYPE: PR1

ORGANISM: Drosophila melanogaster

US-09-270-767-45019

Query Match 4.9%; Score 450.5; DB 4; Length 334;

Best Local Similarity 35.3%; Pred. No. 1.1e-31;

Matches 108; Conservative 48; Mismatches 103; Indels 47; Gaps 10;

QY 1463 VQDSBE-AHS-----ASSFFAETISQYTSIGCDIPGWLWVSLKNGITPYLRCA 1513

DB 51 VESEENQGHGTVPICESHILALLEY-----VQRMSSFLRCS 93

QY 1514 LFPHYLLGVTPPELHTNSAGBSYLSALCSYLSLFTNLFLFQBYWDTV-RPLQRCWADP 1572

DB 94 LFYRFLTDVDPPTDQPD-RFDLMQYGLDPLMGVYFD--METVYATMHSFASHP 150

QY 1573 ALLANCLKOKTVVYPRKNSL-IE-----LPDYSCLLNQASHFRCPRSADDE 1620

DB 151 HIDREVEQR-----CQPDARRSLQVPCPLRPLKVLKVDLINSVSDIFCPNNREE 206

QY 1621 RHPVLCLFCGALCSQNICCOBIVNGEEVGACIFHALHCGAGVCIFPKIRECRVVLVEG 1680

DB 207 MKTPTWCLICGLLCQSYCCQPELKVSVGACTHEAHAGAEVGFIRDRDQVYVL-G 265

QY 1681 KARGCAYPAPYLDEYGETDPLKRGNPLHLSRERYRKLHLVMOQHCHIEIARSQETNOM 1740

DB 266 RGKGCVPVPYLDYGETDGLRGNPLRLSQAARYKIYQLWLHGLHGLAEIARLDNANV 325

QY 1741 LFGFNW 1746

DB 326 AAAAQW 331

#### RESULT 8

US-09-248-796A-19402

Sequence 19402, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 19402

LENGTH: 817

TYPE: PR1

ORGANISM: Candida albicans

US-09-248-796A-19402

Query Match 3.4%; Score 317.5; DB 4; Length 817;

Best Local Similarity 22.2%; Pred. No. 5e-19;

Matches 205; Conservative 125; Mismatches 327; Indels 268; Gaps 42;

QY 961 KLKGPQLRGQDMITWILQMPDTVKRLKSKLIVATTSGSEIKNDSEITHDKKAERK 1020

DB 25 KRSIFKINQK-----YSPFPTLKARSEFDISKVDLSDDANSEDLGKKRIAKER 79

QY 1021 RKAAARLHRQKIMQMSALQRNIETHKLMYDNTSEMPGKEDSIMEESTPAVSDYSRI 1080

DB 80 QAKLMAKPKKQ---QSLFLKKNQPET-----I 103

QY 1081 ALGPKRGPSVTEKEVLT-----CILQBEQVEKIENNAMVLSACVOKSTALTQH 1129

DB 104 ACG-----SDTEMBELDDHAHWKPFPEHHCMLCQNAE---DAGPFGIITAYISKS---CBF 152

#### RESULT 9

US-09-248-796A-19403

Sequence 19403, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA

FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 19403

LENGTH: 631

TYPE: PR1

QY 1130 RQKP-----IELSGEALDPLFMDPDLAYGVTVG----- 1157

DB 153 RNVPFSDRYWFLKAFSDNVDLNGNBYE-LDINGLSNPHSTQTNWNSFMOQASNSSVIGPG 211

QY 1158 -----SCGHVHACVCKQKPYEAVOLSSQQRHVDLPLESGEYLCPCKSL 1203

DB 212 FKLNEFIESHFVALSCGHMHCYQYVNYISNRRNQOITRSTPENMDRKEFLCPCKL 271

QY 1204 CNTVIPILQPKINSENADALAQLLTLARITQTVLARISG-----NRIHAKGENPI 1257

DB 272 KNIFVPIL-----NSSNN-----RSLOEFLKPYAGTSPFLDLDMKIAKORGWT 314

QY 1258 PIFPNQMGDSTLEPHSILSPGVESIKYSIKEMVILPATIYIIGLKVPDRDPRV 1317

DB 315 ERP--QSPSEQBFSRSILT-GIS-----KEMISMDLTK-----ELTKQQRDFRL 356

QY 1318 -----PMLTWSTCAP-----TIOAIENLL---GDEGKPLFGALQNRHGLK 1356

DB 357 MLSDFMQTLTLFTFPQVFKADSGFVLIDTIKISISLRGVSSGGQIILHOLSNILLTLR 416

QY 1357 ALMQF-----AVAQRIITCPQVLIQKHLVRLSVLPNPKIKSEDTPLCLLSIDLHFLVAGA 1409

DB 417 TLNEFKNSSTLMKVNWIDLPN-LRKSAYSKIPIAYIFALSKKSINNYMLEVDFEHLVNT 475

QY 1410 VLAPSLYWDPDVLDQPSVSSSYNHLYLFLHITWAHMLQILL-----TVD 1455

DB 476 D-TFPLSGP-----SFNSVLETCPIGHLI---QSLSIMVNEMSANCLKQNWAYTVS 522

QY 1456 TGLPLAQVQEDSEBAHSASSPFABISQYTS-CSIGCDIP---GWYLVSLKNGITPYLR 1511

DB 523 DIPTIADIPQGISEA--AISCFKLSNSNDVRCIGNDESRRKFGVIYSMLVXIATPFLR 580

QY 1512 AALFFHYL-----LGVTPPEELHTNSAEYSALCSYLSL-PTNLPL-LFOE----- 1556

DB 581 AAIYAVQCANIDGIDPAASPETEI-----EADRLCSFLNLPKVCYELRFLPNESSV 633

QY 1557 ---YWDTVRPLQRCADPALNCLKQKNVTVRYPRKRN-----SLIELPDDYSCLLNOA 1608

DB 634 GSVFYDFL-----TFSSALPKANDVLRIKKQVEYPGIVKLVDPRLDPFFTKY 682

QY 1609 SHFRCPERSADDERKHPVLCLFCGAI--LCSQNICQOEIVNGEVEGACIFAL-HCAGVC 1665

DB 683 YLDRHNNPNKKVEDPAICLFCGDDVVDIQKAIGCKB-----GQCTTHYLKBCSNDVG 735

QY 1666 IFPKIRECRVVLVEGKARGCAYPAPVLDYEGTDPCLKRGNPLHLSRERYRKLHLVMOQ 1724

DB 736 IFLLPKD-RSLLHLKNGGTFNAPFLDEHGBIAESKAKALHLMKPYDDFIRNIWLL 794

QY 1725 HCIIEIARSQETNOMLFGFNWQLL 1749

DB 795 HNVQNCIVRSLES--VLDPGGWETL 817



Db 164 -QAQL-TQAQAQPAOSS-----TEME-----PFVMM 188  
Qy 396 YKOLQKEYISDDHRSISITALSVMQPTVPTLARHLIEBQ-----VTSVIT 442  
Db 189 KOQLQEK-----EEFISTLAQLSQTAQEAQAQVVRREKDAFETQVRLHEDELQVLT 242  
Qy 443 ETLLVLPVYLDNRNKNFGYSQDKLGRVYAVICDKYILIS----KPTIWTERRLM-- 496  
Db 243 QADVETEMQOKVULQKLEHEBESLVGR--AQVDDLOQELTAASQNOILSQQQOME 300  
Qy 497 -----QPLEGRSFLKIILTCMGMEEIRROVGHIEVDPDWEAAIAIQMOLKNTLLMP 549  
Db 301 AEHNTLRNVTETEEESKILLEKWELEVAERKLSFH-----NLQEMGHLLQEP 349  
Qy 550 QEWACDEBL--LLVAYKECHKAVMRCSTSF1--SSSKT--VVOGCHSLETYSYRVSDEL 604  
Db 350 EQAGQAQAELESYSALQKHAEMBEKTSILSLQKTGOELQACDALKQNSKLLQDK 409  
Qy 605 VSHLPLSLTLAGHLVLSLGA-----VSRLE----HEP----VSPEDFQVEVLVEYPLRC 652  
Db 410 NEQVQSAQTIQOEDLOQKSKBSIQFLNRLPQOQHETASQTSFPDPVNEGTAQVTEEN 469  
Qy 653 LVLVAQVLAEMWRNGLSISOVFYQDVKCREMYDKDIIMLOI-----GAS 700  
Db 470 IASLQKRVVLENEKCALILSSI--ELEBELKAENEKLSQQITLAEQNRGADREVRSEIS 528  
Qy 701 LMD-PNK-----FLLLVLQRYELABA 720  
Db 529 IVDIANKRSSAESQDVLENTFSQKHKLSVLLLEKAEQEEIAFLKQLQKGAEEA 588  
Qy 721 FNKTISTKQODLIKQY--NTLIEMLQVLIYVGERY----- 755  
Db 589 DHEVL---DQKEMKQMEGIAPIKMKVFLDTGDQFPLMPNEESSLPAVEKEQASTEQ 645  
Qy 756 -----VPGVGNVTKERTVTRTEIHLCTEPPHSA-- 785  
Db 646 STSBEISLNDAGVELSKTQDQDKSLSAVPDIGHQCHQDELE-RLKSQILELELNFHKAQ 704  
Qy 786 --IAKNLPENNETGLENVINKVATPKPGVSGHGV-----ELKDESL---KDFNMYP- 835  
Db 705 EYERKULDEKAKEISNLQLIBE---FKGNADNNSAFTALSEERDQLLSQVKELSMVTE 761  
Qy 836 --YHYSKTOHSAEAMQKGRQENK---DEALPPPPPPFCFAPSQVKNLNCIDIMVY 890  
Db 762 LRAQVKQLEMNLAERORRLDYESTAHNDLLTE-----QIHSLSTEAKSKDVKIEV 814  
Qy 891 LRTVPE-----RAIDTDSNLWTGMLQMAFH 916  
Db 815 LQNELDDVQLQSPQSBSTLRSQSLQNKSEVLEGAERVRHISKVBELSQALSQKLE 874  
Qy 917 ILALG--LLEBK---QOLQKAPEE-----EVTDFPYHKAIRG-----SSAMNIQMLLEK 961  
Db 875 ITKMDQLLEKRDVETLQATIEEKQOQVTEISPSWTERKMWQLNEBESLGVETKLEQ 934  
Qy 962 LKGI-----PQLEGQDMITWILQMDPTVVRUREKSLVATTSSESINDEITHD 1013  
Db 935 LNLLSRAEAKKEQVEEDNEVSGLKQNYDEM-----SPAQISKEBELQHE 980  
Qy 1014 -----KEKAERKKAARLHROKIMQAQMSALQKQPIETHKLYDNTSEMPKESIMEE 1068  
Db 981 FOLLKKEQORXKQALINKEQLQVRSLEB---ELANLQKESKKEIPLSETERGEV 1037  
Qy 1069 ESTPAVSDYSIALGPKRGPSTVEKVLTCILQOE-----EOEVKIENNAWLS- 1117  
Db 1038 EEDKENKEVSE-----KCVTSK-----COHIEIYKQTISEKEVELQHIRKDLER 1082  
Qy 1118 --ACVQKSTALTOHRKGPTELSCREALDPL 1144  
Db 1083 KLAABEQFQALVKQMNQTLQDQNTQIDLL 1111

RESULT 12

US-09-949-016-6468

; Sequence 6468, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6468  
; LENGTH: 1979  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-6468

Query Match 2.0%; Score 184; DB 4; Length 1979;

Best Local Similarity 18.7%; Pred. No. 2.3e-06;

Matches 233; Conservative 195; Mismatches 453; Indels 366; Gaps 53;

Qy 209 BEKLEPPELQREKKERYCVLFDNHHSHSYDHYVLSQRLDCELAELAAQLHTTAIDKEGR 268  
Db 79 EKHASHIQIKQOSTSTANQL-QQKEVEISH-LKARIALQDQLKLQ-----SAA 128  
Qy 269 RAVKAGAYAAQCAKEDIKSHSENVSRHP--LHVEVL-HSEIMAHQKPAIRLGSMWNKIM 325  
Db 129 QVSPSGAGVPATTAS---SSPAYGISHPSPAFHDDMDPGDIISQOEINRLSNEVSRLE 185  
Qy 326 SVSSDFRQIPQACLRBEPDSENPCLISRLMLWDAKLYGARKILHELIFSPFEMEMKYK 385  
Db 186 SEVGHWRHI-AQTSKAQGTDSNDSQSEICKL----- 214  
Qy 386 KLAFMEFVKYKQKQKEYISDDHRSISITALSVMQPTVPTLARHLIEBQNVISVITEL 445  
Db 215 -----QNIILKELQNRQSEI-DDHQHMSVLQNAHQKLTESRRHREELSDY-----EER 264  
Qy 446 LEVLPEYLDNRNKNFGYSQDKLGRVYAVICDKYILISKPTIWTERRMQFLGFRSF 505  
Db 265 IELENLLQOGG---SGVIEDLSKIYEMOKTIQVLQIEK--VESTKOMEQLEDKIKDI 318  
Qy 506 LKILTCMGMEEIRROVGHIEVDPDWEAAIAIQMOLKNTLLMPQEWACDELLLVAYK 565  
Db 319 NKGLSSAENDRDILRREQELNVEK-----RQIM-----E 348  
Qy 566 ECHKAVMRCs---TSFISSTKTVQSCGSHLETYSYRVSDELVSTH--LPLSRTLGLHV 620  
Db 349 ECENLKLSCSKLPQSAVQSDTM-----TEKERILQASASVEVFRLOQALSADBN 399  
Qy 621 RLSRLGAVSRILHEPVSFEDFQVEVLVRYPLRCLVLVAQVAVMWRNGLSISQVIFYQ- 679  
Db 400 EIMRLSSLNQNSLAE-DNLKLMRIEVLKEKSLLSQKEEL-----QMSLLKLANVEV 454  
Qy 680 -----DVKCREMYDKDIIMLOIAGSLMDPNKPL-----LLVLQRYEL-----ABAFNKT 724  
Db 455 IKSTATRDISDSELHD---LRLANLEAKEQELNQSISEKETLIAEIEBELDRQNEATKHM 511  
Qy 725 ISTDQDLIKQYN---TLIEBMLQVLIYVGERVYVGVNVTKEVTRTEIHLICIBPM 781  
Db 512 ILIKDQ-LSKQONEGDSIISKQDL-----NDEKKRV-----HQLEDDKM 551  
Qy 782 PHSALAKNLPENNETGLENVINKVATPKPGVSGHGVBEKSLKQDKNFYHYSKT 841  
Db 552 D---ITKELDVQKEKLIQSEVALNDLHLTKQK-----LEDKVENLVD-----QLNKS 595  
Qy 842 QHS-----KAEMQKGRKQKRNDEALPPPPPPFCFAPSQVKNLNCIDIMMYILRTVFER 897

Db 596 QESVNSIQENLEKHEIRONEE-----LSRIRNEL-----MQ 629

Qy 898 AIDTDSNLWTEGMLQWAFHILALGALLEBKQOLQKAP--EVTDFYHKASRLGSSAMNI 955

Db 630 SLNQDSNSNFKDTLLKEREAEVRNLKQNLSELEQLNENLKVAFDVKMENKVLACEDV 689

Qy 956 QMLLEK-LKGIPOLEGQKDMITWILQMPD-----TVKKLRKSCLIVATTSGSESI 1005

Db 690 RHQLEBCLAGNQLSLEKNTIVETLKMEKGEIEABLCKWAKLLBEANKYKTEIELSNA 749

Qy 1006 KN-----DBITH-----DKEKAERKRAE-----AARLHQKIMAQM 1037

Db 750 RNLTNSALQLEHEHLIKMKQKMEIAELKKNIEQMDTDHKETKDVLSSELEQKLTOL 809

Qy 1038 SALQKNPIETHKLMYDNTSEMPGK-----EDSIMEES 1070

Db 810 INKKEIFIEKLRSSKLOEELDKYSQALRKNELRQTIEBKDRSLGSMKEENHLOEEL 869

Qy 1071 TPAVSDYSRIA--LGPKRGPVTE--KEVLTCLCQEOEVKIENNAMVLSACVOKSTAL 1126

Db 870 ERUREQSTAPVADPKTLDSTVTELASEVSQLNTIKEHLEBEIKHKQKIEBQNSKQML 929

Qy 1127 TQHRGKPIELSGEALDPLFMDPLAYGTYTSGCGHMHAVCQKYPEAVQLSSQQRHVD 1186

Db 930 LQ-----SLQEQKEMDF-----RY-----QHEQMATHQ 956

Qy 1187 LFDLSGEVLCPLCKSLCMTVIP-----IPLQPKINSENADALAQLLTLARWITVLARI 1243

Db 957 LP-LEKDEBI-----KSLQKTIIEQIKTQLHEERQDQTDSNDFQEB-----TKV 999

Qy 1244 SGYNIRHAKGENPIPIFFNQGMGDSLEPHSILSGVSESIKYSNISIKEMWILFATTIYR 1303

Db 1000 QSLNIENG-----EKHDLKSKATERLVK-----1023

Qy 1304 IGLKVPDPDRPRVPMLTWSTCAFTIQAIENTLLGDEGKPLFGALQNR 1350

Db 1024 -GK-----BRELEIKLNEKNISLTQ-IDLSKDEVGKLTQIIQK 1064

RESULT 13

US-09-949-016-7404

; Sequence 7404, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7404

; LENGTH: 2047

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-7404

Query Match 2.0%; Score 184; DB 4; Length 2047;

Best Local Similarity 18.7%; Pred. No. 2.5e-06;

Matches 233; Conservative 195; Mismatches 453; Indels 366; Gaps 53;

Qy 209 EKELEPPELOIREKNERYCYVLFNDEHSHYDHYVLSQALDCELAELHQTATIDKEGR 268

Db 147 EKHSEASEIQKQOSTSYRNQL--QKEVEISEH-LKARQIALQDQLKIQ-----SAA 196

Qy 269 RAVKAGAYAACQAEKEDIKSHSNVSHQ--LHVEVL-HSEIMAHQKFAELGSMWNKIM 325

Db 197 QSVFSGAGVPATTAS---SSPAYGISHPSAFHDDMDPDGDIISQOEINRLSNEVSRL 253

Qy 326 SYSSDPFOIFQOACLRPEPDSENCLISRLMLWDAKLGARKILHELIPSPFFMEMEYK 385

Db 254 SEVGHWRI--AQTSKAQGTDSQSEICKL-----282

Qy 386 KLFAMEFVKYKQLOKQEYISDHDORSISITALSVQMTPTPTLARHLIEEQNVISVITETL 445

Db 283 ----QNIIKELKQNSQEI--DDHQHMSVLQNAHQKLTESRRHRELSDY-----EER 332

Qy 446 LEVPEYLDNRNKNFQYQSKLGRVYAVICDLKYLIIISKPTIWEHLRQMFLGFRSP 505

Db 333 IEELNLIQOGG---SGVETDLSKIYEMQKTIQVLQIEK--VESTKMEQLEDKIKDI 386

Qy 506 LKILTCMQMEIEIRQVQCHIEVDPDWEAAATAIQMLAKNILLMPQEWACDBELLVAYK 565

Db 387 NKGLSSAENDRIILREBQELNVEK-----RQIM-----E 416

Qy 566 ECHKAVMRCSS--TSFISSTVWQSCGHSLETYSRVSEDLVSIH--LPLSRTILAGLHV 620

Db 417 ECENLKSCSKLOPSAVKQSDTM-----TEKERILAQSASVEEVRLLQALSDAEN 467

Qy 621 RLSRLGAVSRULHEFVSPDFQVEVLVYPLRCLVLVAQVVAEMMRNGLSLISQVYYQ- 679

Db 468 EIMRLSSLNQDNSLAE-DNLKLMRIEVLKESKLSLSQKEEL---QMSLLKLNVEYEV 522

Qy 680 -----DYKCEBMYDKDIIMLOIGASIMDPNKF-----LLVLQRYEL---AFAFNT 724

Db 523 TKSTATRDISLDSSELH---LRLNLEAKEBELNOSISEKETLIAIEBELRQNEQATKHM 579

Qy 725 ISTDQDILIKOVN---TLIEBMLQVLIYIGERYVPGVGNVTKEVTMREIHLICISPM 781

Db 580 ILIKQO--LSKQONEGDSIIISKQDL-----NDEKKEV-----HQLEDDKM 619

Qy 782 PHSAIAKNLPENENNETGLENVINKVATFKPGVSGHGYELKDBSLKDFNMVYHYKST 841

Db 620 D--ITKELDVQKEKLIQSEVALNDLHLTKQ-----LEDKVENLVD-----QLNKS 663

Qy 842 QHS-----KAEMQKRRKQENKDEALPPPPPEPCFAPSKVINLNCDIMMILATVFER 897

Db 664 QESVNSIQENLEKHEIRONEE-----LSRIRNEL-----MQ 697

Qy 898 AIDTDSNLWTEGMLQWAFHILALGALLEBKQOLQKAP--EVTDFYHKASRLGSSAMNI 955

Db 698 SLNQDSNSNFKDTLLKEREAEVRNLKQNLSELEQLNENLKVAFDVKMENKVLACEDV 757

Qy 956 QMLLEK-LKGIPOLEGQKDMITWILQMPD-----TVKKLRKSCLIVATTSGSESI 1005

Db 758 RHQLEBCLAGNQLSLEKNTIVETLKMEKGEIEABLCKWAKLLBEANKYKTEIELSNA 817

Qy 1006 KN-----DBITH-----DKEKAERKRAE-----AARLHQKIMAQM 1037

Db 818 RNLTNSALQLEHEHLIKMKQKMEIAELKKNIEQMDTDHKETKDVLSSELEQKLTOL 877

Qy 1038 SALQKNPIETHKLMYDNTSEMPGK-----EDSIMEES 1070

Db 878 INKKEIFIEKLRSSKLOEELDKYSQALRKNELRQTIEBKDRSLGSMKEENHLOEEL 937

Qy 1071 TPAVSDYSRIA--LGPKRGPVTE--KEVLTCLCQEOEVKIENNAMVLSACVOKSTAL 1126

Db 938 ERUREQSTAPVADPKTLDSTVTELASEVSQLNTIKEHLEBEIKHKQKIEBQNSKQML 997

Qy 1127 TQHRGKPIELSGEALDPLFMDPLAYGTYTSGCGHMHAVCQKYPEAVQLSSQQRHVD 1186

Db 998 LQ-----SLQEQKEMDF-----RY-----QHEQMATHQ 1024

Qy 1187 LFDLSGEVLCPLCKSLCMTVIP-----IPLQPKINSENADALAQLLTLARWITVLARI 1243

Db 1025 LP-LEKDEBI-----KSLQKTIIEQIKTQLHEERQDQTDSNDFQEB-----TKV 1067

Qy 1244 SGYNIRHAKGENPIPIFFNQGMGDSLEPHSILSGVSESIKYSNISIKEMWILFATTIYR 1303







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 08:02:32 ; Search time 157 Seconds  
(without alignments)  
5704.627 Million cell updates/sec

Title: US-10-758-672A-2

Perfect score: 9224

Sequence: 1 MADEEAGGTERMEISAEPLQ.....EIASRQETNQLFGFWQLL 1749

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:\*

1: uniprot\_sprot;\*

2: uniprot\_trembl;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	2 Q8IWV7	Q8iw7 homo sapien
2	8928.5	96.8	1709	2 Q8IWY6	Q8iwy6 homo sapien
3	8617	93.4	1757	2 Q70481	Q70481 mus musculus
4	4327	46.9	1755	2 Q8IWV8	Q8iww8 homo sapien
5	4280	46.4	1755	2 Q6DIB9	Q6dib9 mus musculus
6	4278	46.4	1755	2 Q8CGW0	Q8cgw0 mus musculus
7	4262	46.2	849	2 Q8BN40	Q8bn40 mus musculus
8	4218.5	45.7	1756	2 Q6WKZ8	Q6wkz8 mus musculus
9	4209	45.6	811	2 Q96JY4	Q96jy4 homo sapien
10	3160	34.3	1275	2 Q15057	Q15057 homo sapien
11	3014.5	32.7	1246	2 Q8U031	Q8u031 mus musculus
12	2887.5	29.1	1109	2 Q8K216	Q8k216 mus musculus
13	2416.5	26.2	1679	2 Q7QSV6	Q7qev6 anopheles g
14	2280.5	24.7	1824	2 Q8SX71	Q8sx71 drosophila
15	2276.5	24.7	1824	2 Q9VX91	Q9vx91 drosophila
16	2248	24.4	446	2 Q792M3	Q792m3 mus musculus
17	2232.5	24.2	861	2 Q8BUL9	Q8bul9 m mus muscu
18	1945.5	21.1	1927	2 P91133	P91133 caenorhabdi
19	1782	19.3	777	2 Q8ROV7	Q8rov7 mus musculus
20	1716	18.6	333	2 Q60708	Q60708 homo sapien
21	1692	18.3	329	2 Q75492	Q75492 homo sapien
22	1301.5	14.1	575	2 Q6ZUD0	Q6zud0 homo sapien
23	1204	13.1	229	2 Q8CSK3	Q8csk3 mus musculus
24	1085	11.8	439	2 Q6P212	Q6p212 homo sapien
25	1038	11.3	1958	1 UBRI SCHPO	O60152 schizosacch
26	1020	11.1	2156	2 Q7SGT6	Q7sgt6 neurospora
27	1005.5	10.9	410	2 Q8R130	Q8r130 mus musculus
28	914	9.9	2052	1 UBRR SCHPO	O13731 schizosacch
29	898	9.7	1926	2 Q6BHH3	Q6bhh3 debaryomyce
30	848.5	9.2	1877	2 Q6CFU4	Q6cfu4 yarrowia li
31	709.5	7.7	1945	1 UBRI_KLULA	O60014 kluyveromyc

32 707 7.7 148 2 Q68DN9  
33 689.5 7.5 1950 1 UBRI\_YEAST  
34 686.5 7.4 1900 2 Q75F74  
35 679.5 7.4 1813 2 Q75F75  
36 661 7.2 1814 2 Q6BHH4  
37 639.5 6.9 1965 2 Q6PRD9  
38 606.5 6.6 1225 2 Q9L295  
39 605.5 6.6 1787 2 Q6FW32  
40 598 6.5 212 2 Q9H578  
41 549.5 6.0 1872 2 Q00763  
42 543.5 5.9 1709 2 Q6CJW9  
43 541 5.9 2010 2 Q19330  
44 471.5 5.1 1869 2 Q86N11  
45 470.5 5.1 1284 2 Q94JL2

#### ALIGNMENTS

RESULT 1

Q8IWV7 PRELIMINARY; PRT; 1749 AA.  
AC Q8IWV7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Ubiquitin ligase E3 alpha-1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RH Han H.Q., Kwak K.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AV061886; AAL32103.1; -  
DR Genew; HGNC:16808; UBRI.  
DR GO; GO:0016874; P:ligase activity; IEA.  
DR GO; GO:0004842; P:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000307; Ribosomal\_S16.  
DR InterPro; IPR003126; Znf\_Nrscogmin.  
DR Pfam; PF02207; zf-UBRI; 1.  
DR SMART; SM00396; Znf UBRI; 1.  
DR PROSITE; PS00732; RIBOSOMAL\_S16; UNKNOWN\_1.  
DR KW  
SQ SEQUENCE 1749 AA; 200209 MW; 3AE0E1A749884971 CRC64;

Query Match 100.0%; Score 9224; DB 2; Length 1749;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MADEEAGGTERMEISAEPLQPTQRLASWDDQVDFYTAFLHLAQLVPEIYFAEMPDLE 60  
Db 1 MADEEAGGTERMEISAEPLQPTQRLASWDDQVDFYTAFLHLAQLVPEIYFAEMPDLE 60  
Qy 61 KQESVQMSIFTPLEWYLFGEEDPDICLEKLKHSAGFQLCGRVFKSGETYSYSCDCAIDPT 120  
Db 61 KQESVQMSIFTPLEWYLFGEEDPDICLEKLKHSAGFQLCGRVFKSGETYSYSCDCAIDPT 120  
Qy 121 CVLCMDCFQDSVHKHNRHYKMTSTGGGFCGDTAEWTKGPPFCVNHPEGRAGTIKNSRC 180  
Db 121 CVLCMDCFQDSVHKHNRHYKMTSTGGGFCGDTAEWTKGPPFCVNHPEGRAGTIKNSRC 180  
Qy 181 PLNEEVIVQARKIPPSVIKYVEMTIWEKEKLPPELQIREKNERYCYLFDNEHSHSYD 240  
Db 181 PLNEEVIVQARKIPPSVIKYVEMTIWEKEKLPPELQIREKNERYCYLFDNEHSHSYD 240  
Qy 241 VVYSQRLADCELAQAQLHTTAIDKEGRAVRAGAYAAQCEAKEDIKSHSENVSQHPLV 300  
Db 241 VVYSQRLADCELAQAQLHTTAIDKEGRAVRAGAYAAQCEAKEDIKSHSENVSQHPLV 300  
Qy 301 EVLHSEIMAHQFALRLGSGMNMKIMSYSDFRQIFQCAQLRERPDSENPCILSRMLWDA 360

Db 301 EVLHSEIMAHQFALRGLGWMKIMSYSSDFQIQFQACLRPEPUSNFCILSRMLMDA 360  
QY 361 KLYKGARKILHELIFSSPFMEYKGLFAMEFVKYKQLOKEYISIDHDSISITALSVO 420  
Db 361 KLYKGARKILHELIFSSPFMEYKGLFAMEFVKYKQLOKEYISIDHDSISITALSVO 420  
QY 421 MFTVPTLARHLIEQNVISVITETTLLEVLPYLDRNNKFNFGYSQDKLGRYAVICDLK 480  
Db 421 MFTVPTLARHLIEQNVISVITETTLLEVLPYLDRNNKFNFGYSQDKLGRYAVICDLK 480  
QY 481 YILISKPTTWRLRMQFLGFRSFLKILTCQGMEEIRROVGOHLEVPDMEAAIAIQM 540  
Db 481 YILISKPTTWRLRMQFLGFRSFLKILTCQGMEEIRROVGOHLEVPDMEAAIAIQM 540  
QY 541 QLNKILLMFQEWACACDEELLVAYKECHKAVMRCSTSFSSKTVVQSCGHSLETKSYRV 600  
Db 541 QLNKILLMFQEWACACDEELLVAYKECHKAVMRCSTSFSSKTVVQSCGHSLETKSYRV 600  
QY 601 SEDLVSIHPLSRTLAGLHVRLSRGAVSRLLHEFVSFEDPOVEVLVEYPLRCLVLVAQVY 660  
Db 601 SEDLVSIHPLSRTLAGLHVRLSRGAVSRLLHEFVSFEDPOVEVLVEYPLRCLVLVAQVY 660  
QY 661 AEMWRNGLSLISQVYQDVKCRBMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEA 720  
Db 661 AEMWRNGLSLISQVYQDVKCRBMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEA 720  
QY 721 FNKTIKDDQDLIKOYNTLIEEMQLVLIYVGERYVPGVGNVTKREVTWREIHLICIEP 780  
Db 721 FNKTIKDDQDLIKOYNTLIEEMQLVLIYVGERYVPGVGNVTKREVTWREIHLICIEP 780  
QY 781 MPHSIAKLVLPENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMFYHYSK 840  
Db 781 MPHSIAKLVLPENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMFYHYSK 840  
QY 841 TOHSAEHQKRRKQENKDEALPPPPPEFCFAPSKVINLNCNDIMVILRTVFPERRAD 900  
Db 841 TOHSAEHQKRRKQENKDEALPPPPPEFCFAPSKVINLNCNDIMVILRTVFPERRAD 900  
QY 901 TDSNLWTEGMLQNAFHILALGLLEEQQLQKAPPEEVTDFYHKASRLGSSAMNIQMLE 960  
Db 901 TDSNLWTEGMLQNAFHILALGLLEEQQLQKAPPEEVTDFYHKASRLGSSAMNIQMLE 960  
QY 961 KLKGIPOLEGQKMDITWILQMPFTVRLREKSLCIIVATTSGSBSIKNDIETHDKKAEK 1020  
Db 961 KLKGIPOLEGQKMDITWILQMPFTVRLREKSLCIIVATTSGSBSIKNDIETHDKKAEK 1020  
QY 1021 RKAEARLHROKIMQMSALQKNPIETHKIWDNTSEMPGKBSIMEESTPAVDSYRI 1080  
Db 1021 RKAEARLHROKIMQMSALQKNPIETHKIWDNTSEMPGKBSIMEESTPAVDSYRI 1080  
QY 1081 ALGPKRGPSVTEKEVLTCILCOBEQEVKIENNAWLVSACVQKSTALTQHRGPIELSGEA 1140  
Db 1081 ALGPKRGPSVTEKEVLTCILCOBEQEVKIENNAWLVSACVQKSTALTQHRGPIELSGEA 1140  
QY 1141 LDPLFMDPDLAYTGTGSCGHVMAVCWKYFRAVOLSSQORIHVDLFDLESGEYLCPLC 1200  
Db 1141 LDPLFMDPDLAYTGTGSCGHVMAVCWKYFRAVOLSSQORIHVDLFDLESGEYLCPLC 1200  
QY 1201 KSLCNTVPIPIIPIQOKINSADALAOQLTLARWITQTLARISGYNIRHAKENPIPIF 1260  
Db 1201 KSLCNTVPIPIIPIQOKINSADALAOQLTLARWITQTLARISGYNIRHAKENPIPIF 1260  
QY 1261 FNQGMGDSITLFEHSILSFGVSESIKYSNSIKEMVILFATTIYRIGLKVPPDERDPRVPM 1320  
Db 1261 FNQGMGDSITLFEHSILSFGVSESIKYSNSIKEMVILFATTIYRIGLKVPPDERDPRVPM 1320  
QY 1321 TWSTCAFTPTQAIENILGDEKPLFGALONRQHNGLKALMQFAVQAQITCQVLIQKHLVR 1380  
Db 1321 TWSTCAFTPTQAIENILGDEKPLFGALONRQHNGLKALMQFAVQAQITCQVLIQKHLVR 1380  
QY 1381 LLSVLPNPKSDETPCLLSIDLPHVLGVAVLAPPSLYWDDPVDLPQSSVSSSYNHLPLH 1440

Db 1381 LLSVLPNPKSDETPCLLSIDLPHVLGVAVLAPPSLYWDDPVDLPQSSVSSSYNHLPLH 1440  
QY 1441 LITMAHMLQILLTVDVTGLPLAQVQSDSEAHSSASSFFAISIQYTSIGSGCDIPGHVLAWS 1500  
Db 1441 LITMAHMLQILLTVDVTGLPLAQVQSDSEAHSSASSFFAISIQYTSIGSGCDIPGHVLAWS 1500  
QY 1501 LKNGITTPYLRCALPFHYLLGVTPPEELHTNSAGEYSALCSYLSLPTNLFLFLFOBYWDT 1560  
Db 1501 LKNGITTPYLRCALPFHYLLGVTPPEELHTNSAGEYSALCSYLSLPTNLFLFLFOBYWDT 1560  
QY 1561 VRPLQRWCADPALLNCLKQKNTVVRYPKRNLSLIELPDDYSCLLNQASHFRCPSADDE 1620  
Db 1561 VRPLQRWCADPALLNCLKQKNTVVRYPKRNLSLIELPDDYSCLLNQASHFRCPSADDE 1620  
QY 1621 RKHPVLCLFCGAILCSQNICCOEIVNGBEVGACIFALHCGAGVCIFLKRRCRVVLVEG 1680  
Db 1621 RKHPVLCLFCGAILCSQNICCOEIVNGBEVGACIFALHCGAGVCIFLKRRCRVVLVEG 1680  
QY 1681 KARGCAYPAPYLDYGETDPGLKRGNPLHLRERVKHLVWQOHCIIIEIARSQETNOM 1740  
Db 1681 KARGCAYPAPYLDYGETDPGLKRGNPLHLRERVKHLVWQOHCIIIEIARSQETNOM 1740  
QY 1741 LFGFNWQLL 1749  
Db 1741 LFGFNWQLL 1749

## RESULT 2

QBIWY6 PRELIMINARY; PRT; 1709 AA.  
AC Q8IYW6;  
DT 01-MAR-2003 (TREMELrel. 23, Created)  
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
DE UBR1 E3a ligase (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22340442; PubMed=12434312;  
RA Dgany O., Avidan N., Delaunay J., Kraenow T., Shalmon L., Shalev H.,  
RA Eidelitz-Markus T., Kapelushnik J., Cattani D., Pariente A.,  
RA Tulliez M., Cretien A., Schischmanoff P.O., Iolascon A., Fibach E.,  
RA Koren A., Rossier J., Le Merrier M., Yaniv I., Zaizov R., Ben-Asher E.,  
RA Olander T., Lancel D., Beckmann J.S., Tamary H.;  
RT "Congenital dyserythropoietic anemia type I is caused by mutations in  
codanin-1";  
RL Am. J. Hum. Genet. 71:1467-1474(2002).  
DR EMBL; AF525401; AAO14997.1; --;  
DR GO; GO:0016874; Filigase activity; IEA.  
DR GO; GO:0004842; P.ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P.ubiquitin cycle; IEA.  
DR InterPro; IPR000307; Ribosomal\_S16.  
DR InterPro; IPR003126; Znf.NrecoGmin.  
DR Pfam; PF02207; zf-UBR1; 1.  
DR SMART; SM00396; ZNF\_UBR1; 1.  
DR PROSITE; PS00732; RIBOSOMAL\_S16; UNKNOWN\_1.  
KW Ligase.  
FT NON\_TER 1 1709  
FT NON\_TER 1709 1709  
SQ SEQUENCE 1709 AA; 195278 MW; B2B17231A1020FID CRC64;

Query Match 96.8%; Score 8928.5; DB 2; Length 1709;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1697; Conservative 1; Mismatches 11; Indels 1; Gaps 1;

QY 18 LQPTQRLASWDDQVDFYTAFLHLAQLVPELYFAEMDPDLKQESVQMSIFTPLEWY 77  
Db 1 LQPTQRLASWDDQVDFYTAFLHLAQLVPELYFAEMDPDLKQESVQMSIFTPLEWY 60  
QY 78 LFGEDPDI CLEKLIKHSQAFQCGRVFKSGRTTYSRCDCAIDPTCVLCMDCFQDSVHKHNR 137

Db 61 LFGSDPDI CLEKUKHSGAFQICGRVFKSGTTSYCRDCAIDPTCVLCMDCFQSVHKNR 120  
Qy 138 YKHHTSTGGFCDCGDTTEAMKTPFCVNHPEPGRAGTIKENSRCPLNEEVIVQARKIPPSV 197  
Db 121 YKHHTSTGGFCDCGDTTEAMKTPFCVNHPEPGRAGTIKENSRCPLNEEVIVQARKIPPSV 180  
Qy 198 IKVVEMTIWEEBKEKLPPELOIREKNERYCYVLFNDEHSHYDRIYLSQALDCELAQAQ 257  
Db 181 IKVVEMTIWEEBKEKLPPELOIREKNERYCYVLFNDEHSHYDRIYLSQALDCELAQAQ 240  
Qy 258 LHATAIDKGRRAVKAGAYAAQCAQEKEDIKSHENVSQHPHVEVLHSEIMAKOKPALRL 317  
Db 241 LHATAIDKGRRAVKAGAYAAQCAQEKEDIKSHENVSQHPHVEVLHSEIMAKOKPALRL 300  
Qy 318 GSWNKMYSYSSDFRQIFQACLRREBPDSENPCILSRMLMDAKLYKGARKILHELIPSS 377  
Db 301 GSWNKMYSYSSDFRQIFQACLRREBPDSENPCILSRMLMDAKLYKGARKILHELIPSS 360  
Qy 378 FFMEMEYKCLPAMEFVKYKQLOKEYISDDHDSISITALSVMQFTVPTLARHLIBEONV 437  
Db 361 FFMEMEYKCLPAMEFVKYKQLOKEYISDDHDSISITALSVMQFTVPTLARHLIBEONV 420  
Qy 438 ISVITETLLVPEYLDNRNKNFQGYSDKLGVRVAVICDLKYILISKPTIINTERLRMQ 497  
Db 421 ISVITETLLVPEYLDNRNKNFQGYSDKLGVRVAVICDLKYILISKPTIINTERLRMQ 480  
Qy 498 FLRGFSFLKILTCMQGMEIRQVQGHIEVDPDWEAAIAIQMLKNILLMPOEWCAQDE 557  
Db 481 FLRGFSFLKILTCMQGMEIRQVQGHIEVDPDWEAAIAIQMLKNILLMPOEWCAQDE 540  
Qy 558 ELLLVAYKECHKAVMRCSTFSISSKTVQSGHSLKTSYRVSSEDLVSLHPLSRFLAG 617  
Db 541 ELLLVAYKECHKAVMRCSTFSISSKTVQSGHSLKTSYRVSSEDLVSLHPLSRFLAG 600  
Qy 618 LHVRLSLRGAVSRHREFVSPEDQVVLVEYPLRCLVLAQVVAEWMRRNGLSLISOVFY 677  
Db 601 LHVRLSLRGAVSRHREFVSPEDQVVLVEYPLRCLVLAQVVAEWMRRNGLSLISOVFY 660  
Qy 678 YQDVKCREEMYDXDIIMLOIGASLMDPNKPLLVQRYELAEAFNKTISTKQODLIKQYN 737  
Db 661 YQDVKCREEMYDXDIIMLOIGASLMDPNKPLLVQRYELAEAFNKTISTKQODLIKQYN 720  
Qy 738 TLIEMLQVLIYIVGRYVPGVGNVTKGVTMREIHLICIRPMPSHAIKNNPENENNE 797  
Db 721 TLIEMLQVLIYIVGRYVPGVGNVTKGVTMREIHLICIRPMPSHAIKNNPENENNE 780  
Qy 798 TGLENVINKVATPKKQVSGHGYVELKDESLKDPNNMFYHYSKTOHSAEHMOKKRRQE 857  
Db 781 TGLENVINKVATPKKQVSGHGYVELKDESLKDPNNMFYHYSKTOHSAEHMOKKRRQE 840  
Qy 858 NKDEALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLMTEGMLQWAFHI 917  
Db 841 NKDEALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLMTEGMLQWAFHI 900  
Qy 918 LALGLLEEKQOLQAPKEEVTPDYHAKASRLGSSANNIOMLEKJLKGIPQLEQKDMITW 977  
Db 901 LALGLLEEKQOLQAPKEEVTPDYHAKASRLGSSANNIOMLEKJLKGIPQLEQKDMITW 960  
Qy 978 ILQMPFTVKRLREKSLCLIVATTSGSRSIKNDEITHDKEKARERKKAERAAHLHROKIMQ 1037  
Db 961 ILQMPFTVKRLREKSLCLIVATTSGSRSIKNDEITHDKEKARERKKAERAAHLHROKIMQ 1020  
Qy 1038 SALQKQFIETHKLMYNTSMPCKEBSIMEESTPAVSDYSRIALGPKGSPVTEKEVL 1097  
Db 1021 SALQKQFIETHKLMYNTSMPCKEBSIMEESTPAVSDYSRIALGPKGSPVTEKEVL 1080  
Qy 1098 CILCOBEQEVKIENNAWLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDLAYGTYTG 1157  
Db 1081 CILCOBEQEVKIENNAWLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDLAYGTYTG 1140  
Qy 1158 SCGHVMHVCWKYFPAVQLSSOORIHVDLFDLESGEYLCPLCKSLCNTVPIPILOPOK 1217

Db 1141 SCGHVMHVCWKYFPAVQLSSOORIHVDLFDLESGEYLCPLCKSLCNTVPIPILOPOK 1200  
Qy 1218 INSENADALAQLTLARWIQTVLARISGVNIRHAKGENPIPIPFNQGMDSTLEPHSILS 1277  
Db 1201 INSENADALAQLTLARWIQTVLARISGVNIRHAKGENPIPIPFNQGMDSTLEPHSILS 1260  
Qy 1278 FGVESISKYSNSIKEMVILLPATYIYRIGLKVPPDERDPRVPMLTWSTCAFTIOAIENLLG 1337  
Db 1261 FGVESISKYSNSIKEMVILLPATYIYRIGLKVPPDERDPRVPMLTWSTCAFTIOAIENLLG 1320  
Qy 1338 DECKPLFGALQNRQHNGLKALMQPAVAQRITCQVLIQKHLVRLLSVVLNPKISEDTPL 1397  
Db 1321 DECKPLFGALQNRQHNGLKALMQPAVAQRITCQVLIQKHLVRLLSVVLNPKISEDTPL 1380  
Qy 1398 LSIDLPHVLVGAVALPSPSYWDDPVDLPQSSVSSSSNNHLYLFLHITMAHMLQILLTVDTG 1457  
Db 1381 LSIDLPHVLVGAVALPSPSYWDDPVDLPQSSVSSSSNNHLYLFLHITMAHMLQILLTVDTG 1440  
Qy 1458 LPLAQVQDESEEAHSSAPFAEISQYTSIGSDIPGWYLVWSLXNGITPYLRCAALFFH 1517  
Db 1441 LPLAQVQDESEEAHSSAPFAEISQYTSIGSDIPGWYLVWSLXNGITPYLRCAALFFH 1500  
Qy 1518 YLLGVTPPELHTNSAEGYSALCSYLSLPTNLFLFOEYWDVTRPILQRCWADPALLNC 1577  
Db 1501 YLLGVTPPELHTNSAEGYSALCSYLSLPTNLFLFOEYWDVTRPILQRCWADPALLNC 1560  
Qy 1578 LKQKNTVVRYPRKRNLSIELPDYSCLLQAASHFRCPRSADDERKHPVLCFCGAILCSQ 1637  
Db 1561 LKQKNTVVRYPRKRNLSIELPDYSCLLQAASHFRCPRSADDERKHPVLCFCGAILCSQ 1620  
Qy 1638 NTCQBIQVNGEYVACIFHALHCGAGVCIFLKIREFRVVLVEGKARGCAYPAPYLDYGE 1697  
Db 1621 NTCQBIQVNGEYVACIFHALHCGAGVCIFLKIREFRVVLVEGKARGCAYPAPYLDYGE 1680  
Qy 1698 TDPGLKRGKPLHLRERYRKLHLVMOQHCI 1727  
Db 1681 TDPGLKRGKPLHLRERYRKLHLVMOQHCI 1709  
RESULT 3  
070481  
ID 070481 PRELIMINARY; PRT; 1757 AA.  
AC 070481;  
DT 01-AUG-1998 (TREMELrel. 07, Created)  
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)  
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
DE Ubiquitin-protein ligase E3-alpha.  
GN Names=Ubr1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98318583; PubMed=96531112; DOI=10.1073/pnas.95.14.7898;  
RA Kwon Y.T., Reiss Y., Fried V.A., Hershko A., Yoon J.K., Gonda D.K.,  
RA Sangan P., Copeland N.G., Jenkins N.A., Varshavsky A.;  
RT "The mouse and human genes encoding the recognition component of the  
RT N-end rule pathway."  
RL Proc. Natl. Acad. Sci. U.S.A. 95:7898-7903(1998).  
DR EMBL; AF061555; AAC40165.1; -.  
DR FIR; T14318; T14318.  
DR MGD; MGI:1277977; Ubr1.  
DR GO; GO:000151; C:ubiquitin ligase complex; IGI.  
DR GO; GO:0005515; F:protein binding; IGI.  
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IGI.  
DR GO; GO:0004512; F:ubiquitin-protein ligase activity; IGI.  
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IGI.  
DR InterPro; IPR000307; Ribosomal S16.  
DR InterPro; IPR003126; Znf\_Nrecogin.  
DR Pfam; PF02207; zf-Ubr1; 1.  
DR SMART; SM00396; Znf\_Ubr1; 1.







AC O6DIB9;  
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
 DE E130209G04Rik protein.  
 GN Name=E130209G04Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McWhan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywicki M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC075642; AAH75642.1; -;  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR InterPro; IPR002345; Lipocalin.  
 DR InterPro; IPR000408; Reg chr condens.  
 DR InterPro; IPR003126; Znf\_nrcognin.  
 DR Pfam; PF02207; zf-UBR1; 1.  
 DR SMART; SM00396; Znf UBR1; 1.  
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN 1.  
 DR PROSITE; PS00626; RCC1 2; UNKNOWN 1.  
 SQ SEQUENCE 1755 AA; 199154 MW; E6D413D674PBBFDA CRC64;  
 Query Match 46.4%; Score 4280; DB 2; Length 1755;  
 Best Local Similarity 46.5%; Pred. No. 2.8e-242;  
 Matches 835; Conservative 342; Mismatches 533; Indels 84; Gaps 28;  
 QY 1 MADE---EAGTTER--WEISAEPLPQTPORLASWQVDPYTAFLHLLAQLVPEIYFAEM 55  
 DB 1 MASEMEPEVQADISLECSAE-----EIAERWLQATDLNRRVYOHLCACVPIYCRGP 54  
 QY 56 DPDLKQESVQMSIFTPLEWLYFGEDPDICLKLKHSQ-AFOLCGRVFKSGETTYSCRD 114  
 DB 55 NPPFQKEDTLAHLILGPMWYICAEADPALGFPKLQANKPSHLGGRVFKVGEPTYSCRD 114  
 QY 115 CAIDPTCVLWDCFPQDSVHKNRYKMTSTGGGFCDCGDTPEAKWTGPPCVNHPGPRAGTI 174  
 DB 115 CAVDPTCVLWECFLGSIHRDHYRMTTSGGGFCDCGDTPEAKWEGPYCQKHLSSEVV 174  
 QY 175 -KENSRCPLNEEIVQARKIPSPVIVKVVEMTWEEKELPPELOIREKNERYCVLEND 233  
 DB 175 EEDDPLVHLSSEDVIARTINIPALMFYAVDILLTWEKESLPEDLEVAKSDTTYCMLFND 234  
 QY 234 EHSYDHVIVYSLQALDCELAELQHTTATIDKEGRVAKGAYAAOAEKEDIKSHSENV 293  
 DB 235 EVHTYEQVIYTLQKAVNCTQKEALGPATTVDRDGRSRYRGDFQYCDQAKTVIRNTRSQ 294

QY 294 SOHPLHVEVLHSEIMAHOKFALRLGSMWNKIMSYSSDFRQIFCOACLRPEEPSENPCILS 353  
 DB 295 TK-PLKVQVMSSVAHAFNGFLKALSGLSVIGSYDGLRRILLCVQLQEGPGENSSLDV 353  
 QY 354 RLMLWDKLYKGARKILHELIPSSPFMEMBYKGLFAMBFVKYKQLOKEIYISDDHRSIS 413  
 DB 354 RLMLNDSKLVKGARSVYHQLFMSLLMDLVKYLKALFPAKYNRQLQDFMEDDHERAVS 413  
 QY 414 ITALSVQMTPTTLARHLIEQNVISVITETLEVLPEYLDNNKFNQGIS---QDKLG 470  
 DB 414 VTALSVQFTPTTLARMLLTETENLMTVIFKAPMDHL-KHRDAQGRFQERYTALQAFKR 472  
 QY 471 RVYAVICDKYTLISKPTITWTERLMOFLGERSFLKILTCMQCKBEETRRQVQGHLEVPD 530  
 DB 473 RVQSHILDLKYLISKPTENSDLAQKFLQGFDAFLBLLKCMQGMNDPITRQVQGHLEMP 532  
 QY 531 DWEAAIAIQMLKNLILMPQEMCADEBELLVAYKECHKANVRCSTSFSSSTKTVQVS-C 589  
 DB 533 EWEAAFTLQMKLTHVISMVQDMCALDEKVLIEAYKKCLAVLTQCHGPTDGEQPTLSIC 592  
 QY 590 GHSLETKSYRVEDLVSILHPLSRTLGLHVLRLGAVSRHLHEFVSPEDFQVEVLVEYP 649  
 DB 593 GHSVETIRYCVSQEKVSIHLPISTRLAGLHVLLSKSEVAYKPELLPLSELSPMLIEHP 652  
 QY 650 LKCLVLVAQVVAEMRRNGLSLSIQVYVYODVKREEMVDKDIIMLOIGASLMDPNKPL 709  
 DB 653 LKCLVLCAQVHAGMRRNGFSLVQIYYHNKCRBFDKDIIMLOIGVSMMDPNHFLM 712  
 QY 710 LVLOYELAEAFN-----KTISTK--DQDLIKQYNLTIEMLQVLIYIVGERVPGVGV 762  
 DB 713 IMLSRFELYQLFSTPDYKGRPSFVTHKDVQVQNNLTIEMLYLIIMLGERFNPQGVQV 772  
 QY 763 TKEEVTMBRIHLICIBPMPHSAIAKNLPENNETGLENVINKVATPKKPGVSGHGYE 822  
 DB 773 AATDEIKREIITHQLSIKPMHSELVKSLEPDEKMGESVETSAHFKPKGLTGRGMYE 832  
 QY 823 LKDESLKDFNMYFYHYKTOHSHKAMKOKRQENKDEALPPPPPPPCPAFSKVINLL 882  
 DB 833 LKPECAKEFNLYFYHFSRAEQKABEQKLRKREKEDTALPPALPPCPPLFASLVNLT 892  
 QY 883 NCDIMMYILRTVFERAIDTDSNLMTGMLQMAFHALGLLEEKOOLQKAPREEV--TPDF 941  
 DB 893 QCDVMLYINGTILQNAVEHSGMSAMSEMLQVRLHLIGMALQEEKHLENAVEGHVQTF 952  
 QY 942 YHKASRLGSSANW---IQMLEKLKGIPOLEGOKDITWILQMFDTVKLREKSLIVAT 998  
 DB 953 TQKISKPGDAPHNPSILAMLETQNAPELEAHKDMIRWLLKMFNAIKKIRE--CSSSP 1010  
 QY 999 TSGSISIKNDEITHDKEAKRKAARLHROKIMAOMSALQKNFIETHKLMYDNTSEM 1058  
 DB 1011 VAAEETIMESRRDKAERKKKABIAERLREKIMAOSEMRHFIDENKELFOOTLEL 1070  
 QY 1059 PGKESIMEESTPAVDYSIALGPKRGSPVTEKEVLTCLCOEQEYKIENNAVLGA 1118  
 DB 1071 DTSASATL--DSSPPVSDAALTALGPAQVPEPRQVFTICLCOEQEYTVSGRAMVLA 1128  
 QY 1119 CVQKSTALTQHRKPIELSGEALDPLFMDPLAYGTYTSCGHHMAVACWQKFEAVQLS 1178  
 DB 1129 FVQRSTVLSKDRTKTI-ADPEKYDPLFMHPDLSCGTHTSCGHHMAHACWQRYFDSVQAK 1187  
 QY 1179 SQ-----RIHVDLPDLESSEVLCPLCKSLCNVTIPIIPQKINSEADALQULTLA 1233  
 DB 1188 EQRQORLRHTS-YDVENGFEFLCPLCECLSNVTIPL--LPPRSILSRRLN-PSDQPDLA 1244  
 QY 1234 RMTQTVLARISGNYI---RHAAGEN-----PIPIFNQMGSDTLFHSILSF 1278  
 DB 1245 QWTRAVTQIKVQVLMRLRKHNAADTSSSEDTEAMNIIPPEGRP-----DFY----- 1292  
 QY 1279 GVESSIKYNSIKEMVILFATYIRGLKVPDPRPRVPLMTWTSTCAFTIQAIENLIGD 1338  
 DB 1293 ---PRNPYSDSIKEMLTTFGTAAKYVGLKHPNEGPRVPIILCWGTCAVTIOSIRLSD 1349



```
QY 1179 SQ-----RIHVDLPDESSEVLCPLCKSLCNTVPIIPLQPKINSENADALAQLLTIA 1233
DB 1188 EQRQORQLRHTS-YDVENGEFLCPLCECLSNVTPIPLL-LPPRSILUSRLN-FSQOPDUA 1244
QY 1234 RWTOTVLARISGNYI---RHAKEN-----PIPIPFNQMGDSITLFFHSILSF 1278
DB 1245 QWTRATVQIKVQVLMRKHNAADTSSSEDTAMNIIPIPEGFRP-----DFY----- 1292
QY 1279 GVSESSIKNSIKEMVILPATTIYRIGLVKPPDERPRVPMLTWMSCAPTIQAIENLLGD 1338
DB 1293 ---PRNPYSDSIKEMLTTFCTAAKYGLKVHPNEGDRVPIILCWGTCAVTIQISIERILSD 1349
QY 1339 EGKPLEGALQNRQHNGLKALMQPAVAQRITCPQVLIQKHLVLLSVLVNPKISEDTPCLL 1398
DB 1350 BEKPVFGPLFCRLDDCLSLTRPAAAHWTVALLPVVQGHFCKLFASLVPSDSYEDLPCLL 1409
QY 1399 SIDPLHVLGAVLAPPSLYWDDPVDLQPSVSSVSNHLYLFLHTMAHMLQIILL---TVD 1455
DB 1410 DIDMFHLLVGLVAFPALQCOD-----FGSSSLATGDLHIFHLVTMAHIVQIILLTSCTEE 1464
QY 1456 TGLPLAQVQDSSEAHASGFPABISQYTSGSGIDIPGWYLVSLKNGITPVLRCALAF 1515
DB 1465 NGMD--QENPTGEELAILSLHKTLLHQTGTSALKZAPSGWHLWRSVRAAIMPFKCSALF 1522
QY 1516 FHYLLGVTTPPEELHTNSABGEYSALCSYLSLPTNLLFLFOEYWDVTRPILQRCADPALL 1575
DB 1523 FHYLVNGVAPPDILQV-SGTSHFELCNLYSLPTNLLHFLFOENSIDWNSLIESQNSVEYK 1581
QY 1576 NCLKQKNTVVRYPKNSLIELPDDYSCLLNQASHFRCPSADDERKHPVLCLFCGAILC 1635
DB 1582 RYLNGERGAISYPRGANKLIDLPEDYSSLNQASHFRCPSADDERKHPVLCLFCGAILC 1641
QY 1636 SQNICQEIIVNGEVGACIFHALHCGAGVCFILKIRECRVVLVEGKARGCAYPAPYLDEY 1695
DB 1642 SQSYCCQAELEGEDVGACTAHTYSCGSGAGIFLVRVRECQVFLAGTKGCFYSPPLYDDY 1701
QY 1696 GETDPLGRGNPLHLRERYRKLHLVWOCHIEETARQETNOMLPGFNWOLL 1749
DB 1702 GETDQGLRGNPLHLCOERFRKIKLWQOCHSITEEIGHAQEANTLVGIDMWOHL 1755
```

## RESULT 7

```
QBN40 PRELIMINARY; PRT; 849 AA.
AC QBN40;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Mus musculus activated spleen cDNA, RIKEN full-length enriched
DE library, clone:F830005C07 product:ubiquitin protein ligase E3
DE component n-recogin 1, full insert sequence. (fragment).
GN Names=Ubr1;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
```

Query Match 46.2%; Score 4262; DB 2; Length 849;  
Best Local Similarity 94.3%; Pred. No. 1.1e-241;  
Matches 801; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 1 MADEEAGGTERMBISAEFLPQTQRLASWWDQVDFYATFLHQAQLVPEIYFAEMDPDLE 60  
DB 1 MADEEWDGAERMDVSEPPPLAFQRPASWWDQVDFYATFLHQAQLVPEIYFAEMDPDLE 60  
QY 61 KQESVQMSIFTPLEWYLFEGEDPDICLEKXHSQAPOLCGRVFKSGTTCYSCDCAIDPT 120  
DB 61 KQESVQMSILTPLEWYLFEGEDPDICLEKXHSQAPOLCGRVFKSGTTCYSCDCAIDPT 120

STRAIN=NOD; TISSUE=Activated spleen;  
The FANTOM Consortium,  
RIKEN Genome Exploration Research Group Phase I & II Team;  
"Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs";  
Nature 420:563-573(2002).  
[4]  
SEQUENCE FROM N.A.  
STRAIN=NOD; TISSUE=Activated spleen;  
MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
"Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes";  
Genome Res. 10:1617-1630(2000).  
[5]  
SEQUENCE FROM N.A.  
STRAIN=NOD; TISSUE=Activated spleen;  
MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
Fujisake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
"RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer";  
Genome Res. 10:1757-1771(2000).  
[6]  
SEQUENCE FROM N.A.  
STRAIN=NOD; TISSUE=Activated spleen;  
Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kouda M., Koya S.,  
Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Kura M.,  
Kurihara C., Matsuyama T., Miyazaki R., Murata M., Nakamura M.,  
Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y.,  
Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
Tagawa A., Takahashi F., Takaku-Akai H., Takeda Y., Tanaka T.,  
Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AK089616; BAC40933.1; -;  
DR MGD; MGI:1277977; Ubr1.  
DR GO; GO:0000151; C:ubiquitin ligase complex; IGI.  
DR GO; GO:0005515; F:protein binding; IGI.  
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IGI.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IGI.  
DR GO; GO:000512; F:ubiquitin cycle; IGI.  
DR GO; GO:0006511; F:ubiquitin-dependent protein catabolism; IGI.  
DR InterPro; IPR003126; Znf Nrecogin.  
DR Pfam; PF02207; zf-Ubr1; 1.  
DR SMART; SM00396; Znf UBR1; 1.  
DR PROSITE; PS00732; RIBOSOMAL\_S16; UNKNOWN\_1.  
DR Ligase.  
KW NON TER  
FT NON TER  
SQ SEQUENCE 849 AA; 98231 MW; B957FB7E5D08A89A CRC64;

```

QY 121 CVLQMDCFODSVHKNHRYKMHSTSTGGPCDGDTEAWKTGPFVNVNHPGAGTIGENSR 180
DB 121 CVLQMDCFODSVHKNHRYKMHSTSTGGPCDGDTEAWKTGPFVNVNHPGAGTIGENSR 180
QY 181 PLNEEVIQARRIFPSPVIVEMTWEEBKEKLPPELQIREKNERYCYLVFNDEHSHSYD 240
DB 181 PLNEEVIQARRIFPSPVIVEMTWEEBKEKLPPELQIREKNERYCYLVFNDEHSHSYD 240
QY 241 VIYSQALDCELAELQHTTALDKGRRVAVKAGAAACORAKEDIKSHENVSQHLHV 300
DB 241 VIYSQALDCELAELQHTTALDKGRRVAVKAGAAACORAKEDIKSHENVSQHLHV 300
QY 301 EVLHSEIMAHQKALRLGSMNKMVSYPDFROIQCOACLRPEPDSENCLISRLMLWA 360
DB 301 EVLHSEIMAHQKALRLGSMNKMVSYPDFROIQCOACLRPEPDSENCLISRLMLWA 360
QY 361 KLYKGARKILHELIFSSPFNEMEKYKLFAMEFVKYKQKQKQKQKQKQKQKQKQKQK 420
DB 361 KLYKGARKILHELIFSSPFNEMEKYKLFAMEFVKYKQKQKQKQKQKQKQKQKQKQK 420
QY 421 MFTVPTLARHLIBEONVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK 480
DB 421 MFTVPTLARHLIBEONVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK 480
QY 481 YILISKPTITWELRLMQLFEGFSPFLKILTCMQGMEIRROVQCHIEVDPDWEAAIAOM 540
DB 481 YILISKPTITWELRLMQLFEGFSPFLKILTCMQGMEIRROVQCHIEVDPDWEAAIAOM 540
QY 541 QLNKILLMPQEWACADELLVAYKECHKAVMRCSTSFSSKTVVQSCSHLETYSYRV 600
DB 541 QLNKILLMPQEWACADELLVAYKECHKAVMRCSTSFSSKTVVQSCSHLETYSYRV 600
QY 601 SEDLSIHLPLSRTLAGLHVRLSRLGAVSLHFEVFPDQVEVLVEYPLRLCLVLVAQV 660
DB 601 SEDLSIHLPLSRTLAGLHVRLSRLGAVSLHFEVFPDQVEVLVEYPLRLCLVLVAQV 660
QY 661 AEWRRNGLSISQVYQDVCKREMYDKDIIIMQIGASLMDPNKFLLLVQRYELAA 720
DB 661 AEWRRNGLSISQVYQDVCKREMYDKDIIIMQIGASLMDPNKFLLLVQRYELAA 720
QY 721 FNKTISTKQDILKQNTLIEEMQLVLYTVGRRYVPGVGNVTIKBVTWREIHLICIEP 780
DB 721 FNKTISTKQDILKQNTLIEEMQLVLYTVGRRYVPGVGNVTIKBVTWREIHLICIEP 780
QY 781 MPHSATAKLPENNETGLENVINKVATFKPGVSGHGYBLKOBLSKDFNNYFYHSK 840
DB 781 MPHSATAKLPENNETGLENVINKVATFKPGVSGHGYBLKOBLSKDFNNYFYHSK 840
QY 841 TOHSAKHEM 849
DB 841 TOHSAKHEM 849

```

## RESULT 8

```

Q6WKZ8 PRELIMINARY; PRT; 1756 AA.
AC Q6WKZ8;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Ubiquitin ligase UBR2.
GN Mus musculus;
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22948832; PubMed=14585983;
RX DOI=10.1128/MCB.23.22.8255-8271.2003;
RA Kwon Y.T., Xia Z., An J.Y., Tsaki T., Davydov I.V., Seo J.W.,
RA Sheng J., Xie Y., Varshavsky A.;
RT "Female lethality and apoptosis of spermatocytes in mice lacking the

```

```

RT Ubiquitin ligase of the N-end rule pathway.";
ML. Cell. Biol. 23:8255-8271 (2003).
DR EMBL; AY280958; AAQ17202.1; -.
DR GO; GO:0005515; P:protein binding; IPI.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR000408; Reg chr condens.
DR InterPro; IPR003126; Znf_Nrecogin.
DR Pfam; PF02207; zf-UBR1; 1.
DR SMART; SM00396; Znf_UBR1; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
KW Ligase.
SQ SEQUENCE 1756 AA; 199392 MW; 646991E910776E18 CRC64;

Query Match 45.7%; Score 4218.5; DB 2; Length 1756;
Best Local Similarity 46.1%; Pred. No. 1.2e-238;
Matches 828; Conservative 343; Mismatches 539; Indels 85; Gaps 29;

QY 1 MADE---EAGGTER--MEISAEPLQTPQRLASWDDQVDFYAPLHLAQLVPEIYPAEM 55
DB 1 MASEMEPEVQAIDRSLLCSAE-----BIAGRWLQATDLNREVYQHLAHCVPKIYCRGP 54
QY 56 DPDLKQESVSQMSIPTPLEWYLFGBDPCI CLKLKHSG-APOLCGRVKSGTGYTSCRD 114
DB 55 NPFPLEKEDTLAQHLLGPNWTICADDPALGPKLQANKPSPHLCGRVAVGEPYTSRD 114
QY 115 CAIDPTCVLCMDCFODSVHKNHRYKMHSTSTGGPCDGDTEAWKTGPFVNVNHPGAGT 174
DB 115 CAVDPTCVLCMEFLGSIHRDHQYRMTTWGGGFCDCGTEAWKEGYPYQKHKLSSSEVV 174
QY 175 -KENSRCPLNEEVIQARRIFPSPVIVEMTWEEBKEKLPPELQIREKNERYCYLVFN 233
DB 175 EEDPLVHLSEVIAARTYINPAIFRYAVDILTWEKESLPELSEVABKSDTYTICMLFN 234
QY 234 EHSYDHYVTSIQRALDCELAELQHTTALDKGRRVAVKAGAAACORAKEDIKSHEN 293
DB 235 EHYTHQVYITQKAVNCTQKEAIGPATTVDRGRSVIRYDFQYCDQAKTVLRNISRQ 294
QY 294 SQHPLHVEVLHSEIMAHQKALRLGSMNKMVSYPDFROIQCOACLRPEPDSENCLIS 353
DB 295 TK-PLKQVQVHSSVAAHQNFGLKALSWLGSVIGYSDGLARILLQVGLQSGPDGENSLVD 353
QY 354 RLMLWDKLYKARKILHELIFSSPFNEMEKYKLFAMEFVKYKQKQKQKQKQKQKQKQK 413
DB 354 RLMLANDSKLWKGARSYTHQLFMSLSLMDLYKCLFALRPAKRYERLQRYDVTDDHDEFS 413
QY 414 ITALSQVMTPTPLARHL-IBEQNVISVITETLLEVLPEYLDNRNKNFQGYSG---QDKL 469
DB 414 VADLSVQIFTVFSLAPNAPHRKKNLMTVITKAFMDHL-KHRDAQGRFQFERYTALQAKPF 472
QY 470 GRVYAVICDLKYLISKPTITWELRLMQLFEGFSPFLKILTCMQGMEIRROVQCHIEVD 529
DB 473 REVQSILDLKVLVLSKPTENSDERLQKFLQGFDAFLKCKMGMDPIRQVQGHIEME 532
QY 530 PDWEAAIAIQMLKNILLMPQEWACADELLVAYKECHKAVMRCSTSFSSKTVVQSG- 598
DB 533 PEWEAAFTLQMKLTHVISHVQVQWDCALDEKVLTEAVKCLAVLTQCHGGTDSQPTLSI 592
QY 589 CGHSLKTSYRVSDELVSITLPLSRTLAGLHVRLSRLGAVSLHFEVFPDQVEVLVEY 648
DB 593 CGHSETTRYCYSQBKVSITLPLSRTLAGLHVRLSRLGAVSLHFEVFPDQVEVLVEY 652
QY 649 PLRCLVLVAQVVAEMWRRNGLSISQVYQDVCKREMYDKDIIIMQIGASLMDPNKFL 708
DB 653 PLRCPVLCAQVHAGWRRNGFSVLNQIYYTHNVKRCEREMFDKDIIMLQTVGSMMDNHL 712
QY 709 LVLQRYELAAFN-----KTIETK--DQDLIKQNTLIEEMQLVLYTVGRRYVPGVGN 761
DB 713 MIMLSRFELYQLFSTPDYKGRPSSEVTHQVQVQNTLIEEMLYLTIIMLVGERFNFVGQ 772
QY 762 VTKEEVTWREIHLICIEBPMHSAKLPENNETGLENVINKVATFKPGVSGHGY 821
DB 773 VAATDEIKREITHQLSIRKPMASSELVSLPDKENKETGMESVIESVAHFKKPGLTGRGM 832

```



822 ELKRSKLDPMYFYHYSTQHSKAEHMKRKRKQENKDEALPPPPPPPPPCFAPSKVNL 881  
833 ELKPECAEFNLYFYHFSRAEQSKAEQRKLRKREKEDTALPPPALPPFCFLFASLVNI 892  
882 LNCIDIMYILRTVPERADTDNSLWTEGMLQAFILALGLLEEKQQLQKAPKEEV-TFD 940  
893 LQCDVLYINGTILQWAVEHGSANSESLQRLVHLI GNAEQEKHLENAVEGHVQVFT 952  
941 FYHKASRLGSSAWN---IOMLEKLGIPOLGQKDMITWILQMDFTVTKRLREKSLIVA 997  
953 FTQIKSKPGDAPHNGSPSILAMLETQNSPSEAHKDMIRWLKMFNAIKKIRE--CSSSS 1010  
998 TTSGESIKNDITDTHKEKAEKRYKAEARLHROKIMQMSALQKNPIETHKLMDYNTSE 1057  
1011 PVAAEGTMEESRDKKAERKRKAETARLREKIMQMSQEMQHFIDENKELFOQTLE 1070  
1058 MPKEDSIMEESTPAVDYSIALGPKRGPSTKEVLTICLQEQEQRVKIENAMVLS 1117  
1071 LQTSASATL--DSSPPVSDAULTALGPAGTQVPEPRQFVTCILQEQEVRTVGSRAWLA 1128  
1118 ACVQKSTALTOHRGPIELSGBALDPLFMDPLAYGTYTSGCHVHVAWCQYFEAVOL 1177  
1129 APVQSTVLSKORTKI-ADPEKYDPLFMHPDLSCTHTGSGCHVHVAHCWQRYFDSVQA 1187  
1178 SSQQ-----RIHVDLPDESGLYELCPKLSLNTWVPIPILOPKINSENADALQLLTL 1232  
1188 KBQRQRRLRHTS-YDVENGFLCPCLCELSNTWVPIPL-LPFRSILSRRL-FSDQDPL 1244  
1233 ASWIOQLVARISGYNI---RHAKGEN-----PIPIFFNQGMGDSLTLEPHSILS 1277  
1245 AQWTRAVTQOIKVQMLRKRNAADTSSSEDETEAMNIIPIPGFRP-----DFY---- 1293  
1278 FGVESSIKYSNGIKEMVILFATTIYRIGLKVPPDERPRVPMLTWSTCAFTIOAENLIG 1337  
1294 ----PRNPVSDSIKEMLTFTGTAAYKVLKVPNEGDPVRPILWGTCAFTIOSIERILS 1349  
1338 DEGKPLFGALQNRQHNGLKALMQFAVQRITCPQVLIQHLVRLLSVLPNKSEDTPL 1397  
1350 DBEKPVFGPLPCRLDDCLSLRSLTRFAAAHWTVALLPVVQGHFCKLFASLPSPSYEDLP 1409  
1398 LSTDLPHVLGVAVLAPSLYWDVDPVQLQPSVSSVSNHLYLPHLITMAHMLQILL---TV 1454  
1410 LDIDMFHLLVGLVLAPALQCD-----FSGSLATGDLHIHFLVTHAHVQILLTSCTE 1464  
1455 DTGLPLAQVQDESEEAHSAASSPFAEISQVTSIGCDIPGWYLVSLKNGITPYLRCAAL 1514  
1465 ENGMD--QENPTGEEELAILSLHKLTHQVTSALKEAPSGWHLWRSVRAAIMPLKCSAL 1522  
1515 PPHYLLGVTPPELHTNSAGEYSALCSVLSPTNLFLPFOEYWDVTRPQLQWCAADPAL 1574  
1523 PPHYLVNGVPAPDPLQV-SGTSHPHELCNLYLSLPTNLHLFOENSIMNSLIESWCQNSE 1581  
1575 LNLCKOKTVVRYPRKRNLSILPDDYSCLLNQASHFRCPRSADDERKHPLVLCFCGAIL 1634  
1582 KRYLVNGERGAISYPRGANKLIDLPEDYSSLIQNASNFCPSKGGDSRAPTLCLVCGSLL 1641  
1635 CSQNICQCEVNGEEVAGACIFALHCGAGVCIFLKI RECRVNLVEGKARGCAYPAPYLDE 1694  
1642 CSQSYCCQAELEGEDVACTAHTYSCGSGAGIFLVRVRECVQLFAGKTKGCPSPPYLDD 1701  
1695 YGSTDPLKRGKGNPLHLISRYRYKLHLVWQHCIIIEIARSQETNQMLFGFNWQLL 1749  
1702 YGSTDQGLRGNPLHLQERFRXIKLWQOHSITEIGHAQEANOVLVGDWQHL 1756

RESULT 9  
Q96JY4 PRELIMINARY; PRT; 811 AA.  
AC Q96JY4;  
DT 01-DEC-2001 (T:EMBLrel. 19, Created)  
DT 01-DEC-2001 (T:EMBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (T:EMBLrel. 26, Last annotation update)

Hypothetical protein FLJ14897.  
DE Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
EX PubMed=14702039; DOI=10.1038/ng1285;  
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,  
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
RA Sudo H., Hosono T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,  
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,  
RA Kanehori K., Takahashi-Fuji A., Hara H., Tanase T., Nomura Y.,  
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshioka Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,  
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Nakagawa K.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Yamashita R.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,  
RA "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs";  
RL Nat. Genet. 36:40-45 (2004).  
DR EMBL; AK027803; BAB5380.1; --  
DR InterPro; IPR000307; Ribosomal\_S16.  
DR PROSITE; PS00732; RIBOSOMAL\_S16; UNKOWN\_1.  
FT NON\_TER 811  
SQ SEQUENCE 811 AA; 94337 MW; DA8P2360PFE035P4 CRC64;  
Query Match 45.6%; Score 4209; DB 2; Length 811;  
Best Local Similarity 100.0%; Pred. No. 1.4e-238;  
Matches 811; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 204 MTIWEKEKLPPELOIREKNERYCYVLFNDEHSHYDHYISLQRLADCELAQLHTTAT 263  
DB 1 MTIWEKEKLPPELOIREKNERYCYVLFNDEHSHYDHYISLQRLADCELAQLHTTAT 60  
QY 264 DKEGRAVKAGAYAAQCAKEDIKSHSENVSOHPLFVHLSEIWAHQFALRLGSSWNK 323  
DB 61 DKEGRAVKAGAYAAQCAKEDIKSHSENVSOHPLFVHLSEIWAHQFALRLGSSWNK 120  
QY 324 IMSYSSDFRQIFQCACLREPDSENPCLISRLMLWDKLYKGARKILHELIFSSPFMEWE 383  
DB 121 IMSYSSDFRQIFQCACLREPDSENPCLISRLMLWDKLYKGARKILHELIFSSPFMEWE 180  
QY 384 YKYLFAHEPVKYKQLOKEYISDDHDSISITALSVMQFTVPTLARHLLEQNIVSIVTE 443  
DB 181 YKYLFAHEPVKYKQLOKEYISDDHDSISITALSVMQFTVPTLARHLLEQNIVSIVTE 240  
QY 444 TLLEVLPEYLDNRNKNFNGYQSDKLGRYVAVICDLKYILISKPTTWTERLRMQFLEGFR 503  
DB 241 TLLEVLPEYLDNRNKNFNGYQSDKLGRYVAVICDLKYILISKPTTWTERLRMQFLEGFR 300  
QY 504 SPFKILTCMQGMEETRRQVGHIEVDPDWEAAIAIQMLKILLMPOEWCACDEELLVA 563  
DB 301 SPFKILTCMQGMEETRRQVGHIEVDPDWEAAIAIQMLKILLMPOEWCACDEELLVA 360



QY 564 YKECHKAVMRCSTSPISSTKVVQSCGHSLETSKYRVSBDLSIHLPLSRTLAGLHVRLS 623  
 DB 361 YKECHKAVMRCSTSPISSTKVVQSCGHSLETSKYRVSBDLSIHLPLSRTLAGLHVRLS 420  
 QY 624 RLGAVERSLHEFVSFEDFQVEVLVEYPLRCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 683  
 DB 421 RLGAVERSLHEFVSFEDFQVEVLVEYPLRCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 480  
 QY 684 REBMYDKDIIMLQIGASLMDPNKFLLLVLRQYELAEAFNKTIISTKODLIKQNTLIBEM 743  
 DB 481 REBMYDKDIIMLQIGASLMDPNKFLLLVLRQYELAEAFNKTIISTKODLIKQNTLIBEM 540  
 QY 744 LQVLIIVGERYVPGVNTKEBVTWREIHLICIEPMPHSAIAKPLPENNETGLENV 803  
 DB 541 LQVLIIVGERYVPGVNTKEBVTWREIHLICIEPMPHSAIAKPLPENNETGLENV 600  
 QY 804 INKVAIFKFGVSGHGYELKDBSLKDFNNYFYHYSKTOHSAKEMKRRKQENKDEAL 863  
 DB 601 INKVAIFKFGVSGHGYELKDBSLKDFNNYFYHYSKTOHSAKEMKRRKQENKDEAL 660  
 QY 864 PPPPPPCFAPSKVNLNCDIMMYILRTVPERAIDTDSNLWTEGMLQWAPHILALGLL 923  
 DB 661 PPPPPPCFAPSKVNLNCDIMMYILRTVPERAIDTDSNLWTEGMLQWAPHILALGLL 720  
 QY 924 BEKQOLQKAPEREVTFDHYKASRLGSSANNIOMLEKLGIPLEGOKDMITWILQMF 983  
 DB 721 BEKQOLQKAPEREVTFDHYKASRLGSSANNIOMLEKLGIPLEGOKDMITWILQMF 780  
 QY 984 TVKRLREKSLIVATTSGSESIKNDRIITHDK 1014  
 DB 781 TVKRLREKSLIVATTSGSESIKNDRIITHDK 811

RESULT 10  
 O15057  
 ID O15057 PRELIMINARY; PRT; 1275 AA.  
 AC O15057;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE KIAA0349 protein (Fragment).  
 GN Name-KIAA0349;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97349984; PubMed=9205841;  
 RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,  
 Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. VII.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro."  
 RL DNA Res. 4:141-150(1997).  
 DR EMBL; AB02347; BAA20806.1;  
 DR GO; GO:000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR InterPro; IPR000408; Reg chr condens.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR SMART; SM00184; RING\_1.  
 DR PROSITE; PS00626; RCCL\_2; UNKNOWN\_1.  
 FT NON\_TER 1  
 SQ SEQUENCE 1275 AA; 145331 MW; 101FF1F66E056066 CRC64;

Query Match 34.3%; Score 3160; DB 2; Length 1275;  
 Best Local Similarity 47.3%; Pred. NO. 1.2e-176;  
 Matches 612; Conservative 248; Mismatches 393; Indels 40; Gaps 19;

QY 479 LKYLISKPTIINTERLWQPLEGRSFLKILTCMQMERIRROVGOHIEVDPDWEAAIAI 538  
 DB 1 LKYLISKPTIINTERLWQPLEGRSFLKILTCMQMERIRROVGOHIEVDPDWEAAIAI 60  
 QY 539 QMQLKNILMFQEWACDEBELLVAYKECHKAVMRCSTSPISSTKVVQSCGHSLETSKY 597  
 DB 61 QMQLKNILMFQEWACDEBELLVAYKECHKAVMRCSTSPISSTKVVQSCGHSLETSKY 120  
 QY 598 YRVSBDLSIHLPLSRTLAGLHVRLSRLGAVSRHLHFVSFEDFQVEVLVEYPLRCLVLVA 657  
 DB 121 YRVSBDLSIHLPLSRTLAGLHVRLSRLGAVSRHLHFVSFEDFQVEVLVEYPLRCLVLVA 180  
 QY 658 QVVAEMWRRNGLSLSQVFFYQDVVCRKEEMWYKDIIMLQIGASLMDPNKFLLLVLRQYEL 717  
 DB 181 QVVAEMWRRNGLSLSQVFFYQDVVCRKEEMWYKDIIMLQIGASLMDPNKFLLLVLRQYEL 240  
 QY 718 ABAFN-----KTIISTYK--DQDLIKQNTLIBEMLQVLIIVGERYVPGVNTKEBVTW 770  
 DB 241 YQIFSTPDYGRKPSSEITHKDVVQNNTLIEEMLYILIMLVGERFSPGVQVWATDEIKR 300  
 QY 771 EIIHLICIEPMPHSAIAKPLPENNETGLENVINKVATFKKPGVSGHGYELKDBSLK 830  
 DB 301 EIIHLICIEPMPHSAIAKPLPENNETGLENVINKVATFKKPGVSGHGYELKDBSLK 360  
 QY 831 FNNYFYHYSKTOHSAKEMKRRKQENKDEALPPPPPPPCFAPSKVNLNCDIMMYI 890  
 DB 361 FNNYFYHYSKTOHSAKEMKRRKQENKDEALPPPPPPPCFAPSKVNLNCDIMMYI 420  
 QY 891 LRTVPERAIDTDSNLWTEGMLQWAPHILALGLLEKQOLQKAPERE-VTFDFVHKASRLG 949  
 DB 421 LRTVPERAIDTDSNLWTEGMLQWAPHILALGLLEKQOLQKAPERE-VTFDFVHKASRLG 480  
 QY 950 SSAMN---IQMLLEKLGIPLEGOKDMITWILQMFDTVKRLREKS-CLIVATTSGSESI 1005  
 DB 481 SSAMN---IQMLLEKLGIPLEGOKDMITWILQMFDTVKRLREKS-CLIVATTSGSESI 539  
 QY 1006 KNDRIITHDKKAEKRRKKAABARLHQKIMQAQNSALQKFIETHKLMYNTSEMPGKEDSI 1065  
 DB 540 KNDRIITHDKKAEKRRKKAABARLHQKIMQAQNSALQKFIETHKLMYNTSEMPGKEDSI 597  
 QY 1066 MEESSTPAVDYSRIALGPKRPSVTEKEVLTCILCOBQEVKEIENNAVLACVQKSTA 1125  
 DB 598 MEESSTPAVDYSRIALGPKRPSVTEKEVLTCILCOBQEVKEIENNAVLACVQKSTA 655  
 QY 1126 LTRHGKPIELSGEALDPLFMDPLAYGTYTSCGHVWMAVCWQKYFEAVQLSSQ---- 1181  
 DB 656 LTRHGKPIELSGEALDPLFMDPLAYGTYTSCGHVWMAVCWQKYFEAVQLSSQ---- 714  
 QY 1182 -RIHVDLPDLESGEVLCPLCKSLCNTVPIIPLQPKINSENADALQLTLARVQTVL 1240  
 DB 715 -RIHVDLPDLESGEVLCPLCKSLCNTVPIIPLQPKINSENADALQLTLARVQTVL 771  
 QY 1241 ARISGVNIRHAKGENPIPIFFNQMGDSTLEPHSILSGVESSIKYNSIKEMVILFATT 1300  
 DB 772 ARISGVNIRHAKGENPIPIFFNQMGDSTLEPHSILSGVESSIKYNSIKEMVILFATT 831  
 QY 1301 IYRIGLVKVPDPRPRVPMLTWSTCAFTIQAIENTLLGDEKPLFGALQNRHNGLKALMQ 1360  
 DB 832 IYRIGLVKVPDPRPRVPMLTWSTCAFTIQAIENTLLGDEKPLFGALQNRHNGLKALMQ 891  
 QY 1361 FAVAORITCPQVLIQKHLVRLSVVLVNLKISBDTCLLSIDILFHLVGLAVLAPPSLYWD 1420  
 DB 892 FAVAORITCPQVLIQKHLVRLSVVLVNLKISBDTCLLSIDILFHLVGLAVLAPPSLYWD 951  
 QY 1421 PVDLPQSSVSSSYNHLPLHILITMAHMLQILL--TVDTGLPLAQVQEDSEAHSSSFF 1477  
 DB 952 PVDLPQSSVSSSYNHLPLHILITMAHMLQILL--TVDTGLPLAQVQEDSEAHSSSFF 1004  
 QY 1478 ABISQYTSIGSDIP-GWLVMSLKNGITPYLRCAALFFHYLLGVTPPELHRTNSAEGE 1536  
 DB 1005 ABISQYTSIGSDIP-GWLVMSLKNGITPYLRCAALFFHYLLGVTPPELHRTNSAEGE 1062  
 QY 1537 YSALCSYLSPNLPLFQFYWDVTVRLLQRCADPALLNCLKQKNTVTVYPRKRNLSLIE 1596

```
Db 1063 FEHLCSYSLPNNLCLFPOENSEINSLIESWCRSEVRYLEGEDAIRYPRESKLN 1122
Qy 1597 LPDDYSCLLNQASHFRCPRSADDERKHPVLCIFCGAILCSQNICQOEIVNGEVCACIFH 1656
Db 1123 LPEDYSSLLNQASHFRCPRSADDERKHPVLCIFCGAILCSQNICQOEIVNGEVCACIFH 1182
Qy 1657 ALHCGAGVCIPIKIRECRVVLVEGKARGCAYPAPYLDEYETDGLKRGNPLHLRERYR 1716
Db 1183 TYSCSGVGIFLRVRECVLFLAGTKGCFYSPYLDYGETDQGLRRGNPLHLCKERPK 1242
Qy 1717 KLHLVWQOHCIIIEIARSQETNQLFGFNWQLL 1749
Db 1243 KIQLKWHQHSVTEIEIGHAQBANTLVGIDWQHL 1275

RESULT 11
Q80U31 PRELIMINARY; PRT; 1246 AA.
AC Q80U31;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DE MKIAA0349 protein (Fragment).
GN Name=MKIAA0349;
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT Prediction of the coding sequences of mouse homologues of KIAA gene:
RT cDNAs identified by screening of terminal sequences of 400 mouse KIAA-homologous
RT randomly sampled from size-fractionated libraries."
RL DNA Res. 10:35-48(2003).
DR EMBL; AK122254; BAC65536.3; -.
DR InterPro; IPR000408; Reg. chr. condens.
DR PROSITE; PS00626; RC01_2; UNKNOWN_1.
DR NON_TER 1
SQ SEQUENCE 1246 AA; 140966 MW; C7A0ED7B2D9D265C CRC64;

Query Match 32.7%; Score 3014.5; DB 2; Length 1246;
Best Local Similarity 45.9%; Pred. No. 4e-168;
Matches 599; Conservative 238; Mismatches 374; Indels 95; Gaps 22;

Qy 479 LKYLISKPTIWTIRLMQFLEGFRSFLKILTCMQGMBEIRROVQGHIEVDPDWEAAIAI 538
Db 1 LKVLISKPTWSDLRKQFQDFAPLELLKCMQGMPIRQVQGHIEPEWEAAFTL 60
Qy 539 QMQLKNILMFQWACDEBLLVAYKECHAVRSTSFISSTKVQS-CGHSLETKS 597
Db 61 QMQLKTHVSMVDWCALDEKVLVAYKCKLAVLTQCHGGTDCGEQPTLSICGHSVETIR 120
Qy 598 YRVSDLVSLPLSRVLRLAGLRLGAVSRILHEFVSPEDQVEVLVEYPLRCLVLA 657
Db 121 YCVSEKVSILHPLSRLLAGLHVLKSVAYKFPBELLPLSELSPMLIEHPLRCLVLA 180
Qy 658 QVVAEMWRNGLSLISQVFTYQDYCKREEMDYKDIIIMQLGASLMDPNKPLLLVRLYEL 717
Db 181 QVHAGWRNNGPSLVNQIYYVHNKCRREMFDDKDIIVMLQTVGSMDPNHPLMTLSRFL 240
Qy 718 AEPN-----KTIETK--DODLKYNTLIEMLQVLIYVGERYVGVGNVTKKEVYMR 770
Db 241 YQLFSTPDYGRFSSEVTHKDVVQNNLTIEMLYLIIMLVGERFNPFGVQVAATDEIKR 300
Qy 771 EITHLCIEPMPSAIAKNLPENNETGLENVINKVATPKPGVSGHYVELKDESLKD 830
Db 301 EIIHQLSIKPMASSELVKSIPED-----ELKPECAKE 333
```

```
Qy 831 FNMVYHYSTKQHSKAEHMQKRRKQENKORALPPPPPEFCFAPSKVINLLNCDDIMMYI 890
Db 334 FNLFIYHSRAQSAQEAQRLKREKEDTALPPALPPFCPLFASLVNLIQCDVMIYI 393
Qy 891 LRTVFERAIDTDSNLTGMLQMAFHIALGLLBEKQOLQKAPBEEV-TFPFYHKASRLG 949
Db 394 MGTILQWAVEHHGSAWSSMLQVLLHGLMALQEKHLENAVEGHVQVTFPTQKISKPG 453
Qy 950 SSAMN---IOMLEKLGKIPOLGKQDMITWILQFDTVKRLREKSLIVATTSSESITK 1006
Db 454 DAPHNSPILAMLETQNAPELHAKDMIRLLKMFNAIKIRE--CSSSSVAEAGTYI 511
Qy 1007 NDEITHDKEKAEKKEKAEARLHROKIMQASALQKNEFIETHKLMYDNTSEMPGKEDSJM 1066
Db 512 MEESRDNDKAEKCKKABIRLRREKIMQASQEMORHFIENKELFQOTLELDTASATL 571
Qy 1067 EESTPAVDYSRIALGPKRGSVTEKEVLTCLCOEEOEVKIKENAMVLSACVQKSTAL 1126
Db 572 --DSSPPVSDAALTALGPTQVPE--PRQFVTCILCOEEOEVTGSGRAMVLAAPVQRSTVL 627
Qy 1127 TOHRKPIELSGEALDPLFMDPDLAYGTYTSCGHVMAVCHWKYFEAVQLSSQ----- 1181
Db 628 SKDRTKTI-ADPEKYDPLFMHPDLSCGTHTSCGHVMAHCHQRYFDSVQAEORQORL 686
Qy 1182 RTHVDLPDLSEGYLCPCKSLCNTVPIPILOPOKINSENADALAQLTLARWITQVLA 1241
Db 687 RUHTS-YDVENGELFCLCECLSNVTIFEL-LPPRSILSRRLN-FSDQPDLAQWTRAVTQ 743
Qy 1242 RISGYNI---RHAKGEN-----PIPIFNQMGDSTLPHSLISFGVSSIKYI 1286
Db 744 QIKVVQMLRRKHNAADTSSEDETEAMNIIPIEGFRP-----DFY-----PRNPY 788
Qy 1287 SNSIKEMWILPATTIYRIGLVKPPDERPRVPMLTWSTCAFTIOAIENLGLDEGLECA 1346
Db 789 SDSIKEMLTTFGTAAKYKGLKVPHEGDPVILCMGTCAFTIOIERILSDEKRPVGP 848
Qy 1347 LQNRQHNGIKALMQFAVQRITCPQVLIQKHLVLLSVLPNIKSEDTPCLLSIDLPHVL 1406
Db 849 LPCRDDCLRSUTRFAAAHNTVALLPVVQGHFCKLFASLVPSDSVEDLPCILIDMFHLL 908
Qy 1407 VGAVLAFPSLYWDDFVDLPQSSVSSYNHLYLFLHITWAHMLQILL---TVDTGLPLAQV 1463
Db 909 VGLVLAFFALQCOD-----FSGSSLATGDLHIFLVTMAHIVQIILLTSCTEENGMD--QE 961
Qy 1464 QDSSEAHASAPFAISQYTSIGCDIPGHVYLVSLKNGITPYLRCAALFHYLLGYT 1523
Db 962 NPTGEEELAILSLHKLTHQYTGSAKLEAPSGHMLWRSVRAAIMPFLKCSALPHYLVGP 1021
Qy 1524 PPEELHTNSAEGEYSALCSYLSLPTNLPLLPQYVDTVRPLLQRCWADPALLNCLKQKNT 1583
Db 1022 APDDIQQV-SGTHFHELCNLSLPTNLHLFQENSDIMNSLIESWQCNSEVKRYLNGERG 1080
Qy 1584 VVRYPRKRNLSIELPDDYSCLLNQASHFRCPRSADDERKHPVLCIFCGAILCSQNICQOE 1643
Db 1081 AISYPRGANKLIDLPEYSSLLNQASHFRCPRSADDERKHPVLCIFCGAILCSQNICQOE 1140
Qy 1644 INGEVGVACIFHALHCGAGVCIPIKIRECRVVLVEGKARGCAYPAPYLDEYETDGLK 1703
Db 1141 ELEGEDVGACTAHTYTSYSCSGAGIFLRVRECVLFLAGTKGCFYSPYLDYGETDQGLR 1200
Qy 1704 RGNPLHLRERYKRLHVLWQOHCIIIEIARSQETNQLFGFNWQLL 1749
Db 1201 RGNPLHLRERYKRLHVLWQOHCIIIEIARSQETNQLFGFNWQLL 1246
```

## RESULT 12

```
Q8K216 PRELIMINARY; PRT; 1109 AA.
ID Q8K216;
AC Q8K216;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
```

```

DE E130209G04Rik protein.
GN Names=E130209G04Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skaleka U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RW [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031403; AAH31403.1; -
DR MGD; MG1:1925978; E130209G04Rik.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR004008; Reg chr condens.
DR SMART; SM001841; Znf_ring.
DR PROSITE; PS00626; RCC1.2; UNKNOWN.1.
SQ SEQUENCE 1109 AA; 12528 MW; B506F24067283279 CRC64;

Query Match 29.1%; Score 2687.5; DB 2; Length 1109;
Best Local Similarity 46.8%; Pred. No. 5.5e-149;
Matches 533; Conservative 209; Mismatches 333; Indels 65; Gaps 19;

QY 644 VLVEYPLRLCLVLAQVVAEWRNGLSLISQVYQDVQKREBYDKDIIIMLQIGASLMD 703
DB 1 MLIEHPLRLCLVLAQVHAGWRNGLSLVNIQYIYHNVKREMFDKDVMVLQTVGSMND 60

QY 704 PKFKLLVLQRYELAEAFN-----KTIATK--DODLIKOYNTLIREMLQVLIYGERYV 756
DB 61 PNHFLMILSRFELYQLFSTPDYDYGKRFSEVTHKVVQQNNTLIEMLYLIIMLVGERFN 120

QY 757 PGVGNVTKSEVTRREIILLICIEPMPSAIAKLPENNNETGLENNIVATFKKPGVS 816
DB 121 PGVQVAANDIEIKREIHLQSLKPMASSELVSLPEDEKNETGSMESVIESVAHFKEGLT 180

QY 817 GHGVYELKDESUKDFNMYFYHYSKTOHSAKHMQRKQKQKQKQKQKQKQKQKQKQKQ 876
DB 181 GRGVYELKPECAKEFNLYFYHFSRAEQSKAEBAQRKLEKEDTALPPFPALPPFCPLFA 240

QY 877 KVINLNCIMMYILATVVERALDTSNLTWTEKMLQMAFHIALGLLEKEKQLOKAPPE 936
DB 241 SLVNILOQDMVIMYITLQWAVEHHSASWSMLQRVLHILGMALQSEKHLENAVEGH 300

QY 937 V-TFDPYHKASRLGSSAMN---IQMLEKLKGIPOLEGOKDMITWTLOMFDVTKRLREKS 992
DB 301 VOTFTFTQIKSPGDAPHNSPSILAMLETLOVAPSLEAKDMIRWLLKMFNAIKIRE-- 358

QY 993 CLIVATTSGSESINKDEITHDKAEKRAKAEARLHRQKIMAQMSALQKNFIETHKLMY 1052

```

## RESULT 13

```

Q7QEV6 PRELIMINARY; PRT; 1679 AA.
ID Q7QEV6
AC Q7QEV6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bb1P8105 (Fragment).
GN Name=eb1G8105; ORPNames=ENSANG00000006115;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

```

DR	EMBL; AA01008846; EAA06476.1; -	
DR	GO; GO:0004842; Pubiquitin-protein ligase activity; IEA.	
DR	GO; GO:0006512; Ubiquitin cycle; IEA.	
DR	InterPro; IPR003126; Znf NrecoGmin.	
DR	Pfam; PF02207; zf-UBR1; 1.	
FT	NON_TER 1679 1679	
SQ	SEQUENCE 1679 AA; 192719 MW; 2D23CEB4E4237708 CRC64;	
Query Match 26.2%; Score 2416.5; DB 2; Length 1679;		
Best Local Similarity 33.0%; Pred. No. 8.4e-133;		
Matches 593; Conservative 318; Mismatches 648; Indels 237; Gaps 47;		
QY	74 LEWYLPGEEDPDICLEKLEHSGAP-QLCGRVFKSGETTYSCRDCAIDPTCVLWDCFDQDSV 132	
DB	1 MELFICGPNPDNYLEIQADTVSSVCGRVFKIGETPTSCRECSMDPTCVLSTCFKXSE 60	
QY	133 HKNHRYKMTSTGGGCDGCDTEAKTGPFCVNHBPGRAGTIKENSRCPLNEEVIVQARK 192	
DB	61 HKNHRYKMTSTGGGCDGCDTEAKTGPFCVNHBPGRAGTIKENSRCPLNEEVIVQARK 192	
QY	193 IPSVTKYVEMTIWEEKELEPELQIREKERNRYC-VLFNDHHSYDHYVLSQALDC 251	
DB	118 VFRAILNYSQMQLMHGTGNC-PEL---EDDENTHCTILYNDETHTFEQVIQTLSIVKC 173	
QY	252 ELAERQLHTTADKEGRRAVKAGAYAAQOEAKEDIKSHSE---NVSQHPLHVEVLHSEI 307	
DB	174 EHKALEYVTSIDREGRVAVKASFEVCKLXEDIKENKAMRSTLTTRAMPKVTVMHNE 233	
QY	308 MAHQFALRLGSMWKNIMSYSDFRQIFQACILREBPDSNPCLISRLMDAKLYKGAR 367	
DB	234 VACOHAMQLLVMLQEFLENSTFRSFAETISKPOVTYNLKFILSN---DHNLWKSAR 289	
QY	368 KILHELIFSGPFMEYKYLPAWEFVKYKQLOKEVYISDDHRSISITALSVMQFTVPTL 427	
DB	290 ACWHRLLISGMLMEYENKLLATFTFKLYTSLMQDFIRDDHYHSFISIVSLVQLFTVPTI 349	
QY	428 ARHLIEEQNVISVITELL-EVLPEYL-DRNNKFNFGVSDQKLGKRVAVAVICDLKYLIS 485	
DB	350 AHYLIEENSAFFKLMHTYSETIDKYVKNRQLVFIKNTSTMSFKAAVILIDLKYLISF 409	
QY	486 KPTIWTERRMQPLEGRFPLKILTQWGMEEIRROVGHIEVDPDWEAAIAQMLKNI 545	
DB	410 KEKWTPELVGFLHGMQLIRLLKYMQGDAAATROVGHLEVEQEWETAFTLHLKLSL 469	
QY	546 LLMFQWCADEBELLVAYKECHAVMRCSTSPIS-SKTTVV-QSCGHSLETKSYRVSDE 603	
DB	470 ITLVLEWCATDRMLVLRWVITAL--TETKFALESETVVCCEAANHSASCLTYDVLTR 527	
QY	604 LYSIHLPLSRTLAGLHVRLSGLAVSRHLHFEVSPEDFQ-----VEVLVEYPLRCLVLV 656	
DB	528 PVSVHLPLRFLAGLVTVFER-----HDFTPKVQSOMLPDRPTPEQIEPVLCAITWM 580	
QY	657 AQVVAEMWRNGLSLISQVFPYQDVKCRBEMVDKQIIMLQIGASLMDPNKFLLLVLQRYE 716	
DB	581 SQVQAGWRNRYGALLINQLPFFYRNVKRYEMLRDILVILQMGASLIEANEFILHVLNKKY 640	
QY	717 LAEAFNKTI-----STKQDLIKQNTLIEMLQVLIYVGERVYVPGVGNVTKEE 766	
DB	641 LLGLWLEQDVPQRQAVDAGSGSDEDMRQAGILVBEFLLELLIIVIGERYVPGVGDVTESD 700	
QY	767 VTMREIHLICIEPMPHSAIAKNLPENENNETGLENNVINKVATFKKP-GVSGHGVYELKD 825	
DB	701 RIYKKEIVQVCIKPHSHSELSRALNEDSCSEINFESVIDEAVFEPKNSAKKGVVILKP 760	
QY	826 ESKLDFMVPYHYSKTQHSKAEHMQKRRQENKDEALPPPPPEFCFAPSKVINLINC 885	
DB	761 EYNSWNLYHYHSEKDESEERQRLQKEKNELVCCPPPTPLKLTQIFSTPTILQCD 820	
QY	886 IMMYILRTVERAIDTDSNLWTEGMLQMAFHIALGLLEKEQLOKAPREEVTF----- 939	
DB	821 VMLKVHVVNTRALDLRSVSTEGQLQV-----QTHVTFRTVKKQ 861	
QY	940 --DPYHKASRLSGSSAMNIQMLEKLGIPQLEGQKDMITWILQMFDTVVKLRKSCLIVA 997	

DB	862 ANDNYSRTPFESSGLAQSRSYILCIQFPFVESLRDLRLRWIORYKAMEAKRISS--IQS 919	
QY	998 TTSGESIKNDBITHDKEKAERKABAAARLHQIKIMQMSALQKNFIETHKLAY-----D 1053	
DB	920 EIDRGESSGGLAMADLDLEKEQAKLAARRAQALMAQMQAOQQTFTMTNDFMSTESD 979	
QY	1054 NTEMPGKEDSIMEESTPAVDYSRIALGPKR---GPSVTEKSVLTLCICQEEQEVKIE 1110	
DB	980 DSSEL-GQTEPEQSDMDWQTSVHLVALGPNRTVPKPADTAAALRHSCILCSEKSVLGYT 1038	
QY	1111 NNA-MVLACVQKSTAL-----TQHRGKPIELSGEALDPLFMDPLAYGYTYTSGCHVMH 1164	
DB	1039 HSNVMAVAAFPQSSSVLSRYQQTDERGQ-LQLIETKTHP---SPHLT-----TCGHVMH 1088	
QY	1165 AVCWQKYFEAVOLSSQQRHVD-----LPDLESGEYLCELCCKSLCNVTVPILPQOKINS 1220	
DB	1089 TTCFDPKYNNEIKENRRPYRNPAPILFDIEKNBEFLCLCRFLSNCLLPVPHETITSS 1148	
QY	1221 ENA--DALAQILTL-----ARWIQTVLARISGYNIHAKGNPIPIFN----- 1262	
DB	1149 AKARTNASASLTSTDEEMDSDEVRWM-----NP-PIHLNTLVLRDV 1190	
QY	1263 QCMGDSITLFEHSILSFGVSESIKYSIKEMVILFATTIYRIGLVPPDERDPRVPLTW 1322	
DB	1191 BELHGVRLRYGYTTHDPPPEISRLPDRPIKYSILNFAKNV-RSFTTVPYKAKFAEYLA 1249	
QY	1323 STCACTIAIENMLLDEGKPLFGALONQHRNGLKALMQPAVAQRITCPQVLQKHLVRL 1382	
DB	1250 LTLSTYIKSIEMQLRMQRPGLSGELSIYSTCLTG-----LVRAA 1289	
QY	1383 SVVLNPKISEDIT---PCLLSIDLPHVLAGVLAFLPSLYWDDPVDLPQSSVSSYNHLYLF 1439	
DB	1290 SANGPLVKSMTAHYTMLLFLVDVITLFG-LKPSFSFPEWDLFGLLNTCLFTTSTRVLSYF 1348	
QY	1440 -----HLITMA-----HMLQILLTVDTGLPLAQ---VQEDSEBAHSASSFPABI 1480	
DB	1349 VTERMPKGDITLHDSIVQAI FVVNMMLRTIITSVGTGTAQDTPMADGADEGCESYAKQK 1408	
QY	1481 SOYTSGSIGCDIPGNYLWVLSKNGITPY-----LRCAALFFHYLGVTPPEELHNTSASG 1535	
DB	1409 EDSGKGP-----LTDARENLLTLFQHTLLRCCCLFHGVTDIELPQ---CTDVPPEL 1457	
QY	1536 EYSALCSYLSLPTNLFLLPQYWDVTRPLQKWCADPA-----LLNCLKQKNTVRYPR 1589	
DB	1458 DYQPMKYLDEL-----VDPLAYRTGEPTTELVERLARCNADQIERMYAR 1504	
QY	1590 KR-----NSLIELPDDYSLNQAASHFRCPRSADDERKHPVLCPCGAI 1633	
DB	1505 QAQLGEPKSLPQPALPVRQLDLDLDDYSLNLSVSLFTCPNNIRDSRNPMTCLVCGEI 1564	
QY	1634 LCSQNICQEIYNGEBEVGACIFHALHCGAGVCIFLKIIECRVVLVEGKARGCAYAPAYLD 1693	
DB	1565 LCSQFCCQKELDKSPVSGCTVHTAECGAGIGIFLIRDAEILL--GINKKCFIPAPYLD 1623	
QY	1694 EYGETDPLKRCNPLHLSERYKHLVWQOHCIIIESIARSQETNOMLGFENWOLL 1749	
DB	1624 EYGETDQGLRRGNPLRCKERYKHLVWITLHGLHEEITRRNEAQQTIFATQWQL 1679	

RESULT 14

Q8SK71	PRELIMINARY; PRT; 1824 AA.
ID	Q8SK71
AC	Q8SK71;
DT	01-JUN-2002 (TrEMBLrel. 21, Created)
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	LD31957P.
GN	ORFNames=CS9086;
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.





Db 1650 -CQPDARRSLQVVPCLRLPLRLKVLCDPFDLSLNSVSDIFCPNNEREEMKTPMCLICGL 1708  
 Qy 1633 ILCSQNICCOEIVNGEVGACIFALHCGAGVCFILKIBECRWLVGKARGCAYPAPYL 1692  
 Db 1709 ILGQSYCCQPELGKRGVSGACTHHAACGAEGVIFLIRIDCQVYL-GRGKCGFVPPPYL 1767  
 Qy 1693 DEYGETDGLKRGNPLHLGRYRKLHLVQOHCIEETARSQETQVMQLFGFNW 1746  
 Db 1768 DEYGETDMGLRGNPLRLSQAYRKLYLQWLHGLHGEIARLNDANVAQAQW 1821  
 RESULT 15  
 Q9VK91  
 ID Q9VK91 PRELIMINARY; PRT; 1824 AA.  
 AC Q9VK91;  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE CG9086-PA.  
 GN ORFNames=CG9086;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., B.D.,  
 RA Brandon R.C., Rogers Y.H., Blazef R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.P., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foslberg C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler P., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirkas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodgett J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RP MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 SEQUENCE FROM N.A.  
 RP MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 SEQUENCE FROM N.A.  
 RP MEDLINE=22426069; PubMed=12537572;  
 RA Mistra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 SEQUENCE FROM N.A.  
 RP FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 SEQUENCE FROM N.A.  
 RP FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003504; AAF48687.2; -  
 DR FlyBase; Fgmn0030809; CG9086.  
 DR GO; GO:0000151; Cubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; Eubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; Fzinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR InterPro; IPR003126; Znf Nrecoign.  
 DR InterPro; IPR001841; Znf ring.  
 DR Pfam; PF02207; zf-UBR1; I.  
 DR SMART; SM00184; RING1.  
 DR SMART; SM00396; Znf-UBR1; 1.  
 SQ SEQUENCE 1824 AA; 208359 MW; 79992C11175E82A0 CRC64;  
 Query Match 24.7%; Score 2276.5; DB 2; Length 1824;  
 Best Local Similarity 31.7%; Pred. No. 1.6e-124;  
 Matches 607; Conservative 307; Mismatches 707; Indels 293; Gaps 56;  
 Qy 5 EAGGTGMEISAEPLPQTPORLASWMDQVDFYTAFLHLLAQLVPEIYFAEMDPDLKQBE 64  
 Db 29 QAGTLDRSDIIEFLKRESPK---YFYDT---SATVKDTNVTILKCMPE---SLAKEE- 78  
 Qy 65 SVQMSIFTPLEVLGEDPDICLEKLKHSQ-AFOLCGRVKSGTETYSCEDCALDPTCVL 123  
 Db 79 -----IIDVVVEFMLGNDPSSALEKRLRGNTATVCGKVKNGEPTYSCEGVDPTCVL 133  
 Qy 124 CMDCFDSDVHKHRYQWHTSTGCGCDGDTBAWTKGPPCVNHPGPRAGTITKENSRCPLN 183  
 Db 134 CVCNFKSAHRFAKYKMSISGGGCCDCGDDKAWKQYCEHLANRKNPLESKI---LT 190  
 Qy 184 BEVIVQARKIPPSVTKYVVMTEIWEELPELQIREKN-----BRYCVLNDHEH 235  
 Db 191 DAVLERVEICFCAILAFVCVSYL-----EIEPNASLQCLDGNVEGGQVDGAQYCTVLYNDES 246  
 Qy 236 HSYDHYVLSQALCELAEALHTTAIDKEGRVAKGAYAAQCEAKEDIKSH----- 289  
 Db 247 HTFDQVQIOTLTAKIAKRAKADMEIVAAIDREGRAVVKCDTPECNKLVKSIENQMLPTS 306  
 Qy 290 --SENVSQHPHVEVLHSEIMAHQKFAIRLGLSGWMNKIMSYSSDFRQIFCOACLRBBPDSE 347





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 15:31:48 ; Search time 133 Seconds  
(without alignments)  
5080.389 Million cell updates/sec

Title: US-10-758-672A-2

Perfect score: 9224

Sequence: 1 MADEAGGTERMEISAEPLQ.....ETARSGETNQLGFWQLL 1749

Scoring table: BLOSUM62

Gapol 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10F\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	16 US-10-758-672A-2	Sequence 2, Appli
2	9224	100.0	1749	16 US-10-357-819-10	Sequence 10, Appl
3	9224	100.0	1749	17 US-10-758-636A-2	Sequence 2, Appli
4	9168	99.4	1738	14 US-10-287-218-1	Sequence 1, Appli
5	9168	99.4	1738	16 US-10-474-291-1	Sequence 1, Appli
6	9087	98.5	1734	16 US-10-758-672A-19	Sequence 19, Appl
7	9087	98.5	1734	17 US-10-758-636A-19	Sequence 19, Appl
8	8617	93.4	1757	16 US-10-758-672A-15	Sequence 15, Appl
9	8617	93.4	1757	17 US-10-758-636A-15	Sequence 15, Appl
10	4327	46.9	1755	16 US-10-758-672A-4	Sequence 4, Appli
11	4327	46.9	1755	17 US-10-758-636A-4	Sequence 4, Appli

12	4278	46.4	1755	16 US-10-758-672A-6	Sequence 6, Appli
13	4278	46.4	1755	17 US-10-758-636A-6	Sequence 6, Appli
14	3160	34.3	1275	16 US-10-755-889-635	Sequence 635, App
15	3066	33.2	594	15 US-10-094-749-2555	Sequence 2555, Ap
16	2710.5	29.4	1109	9 US-09-529-063-55	Sequence 55, Appl
17	2710.5	29.4	1109	14 US-10-414-378-55	Sequence 55, Appl
18	1169.5	12.7	452	15 US-10-264-237-2544	Sequence 2544, Ap
19	550	6.0	1270	16 US-10-437-963-193003	Sequence 193003,
20	468	5.1	147	15 US-10-140-047-2153	Sequence 2153, Ap
21	453.5	4.9	951	17 US-10-480-988-17	Sequence 17, Appl
22	414.5	4.5	203	15 US-10-264-237-1513	Sequence 1513, Ap
23	386.5	4.2	889	16 US-10-425-115-307970	Sequence 307970,
24	296	3.2	250	15 US-10-286-115-1087	Sequence 1087, Ap
25	295	3.2	247	14 US-10-050-704-1107	Sequence 110, App
26	295	3.2	247	16 US-10-798-512-110	Sequence 110, App
27	276	3.0	81	14 US-10-016-349A-227	Sequence 227, App
28	275	3.0	499	15 US-10-425-114-68139	Sequence 68139, A
29	259	2.8	550	15 US-10-424-599-182689	Sequence 182689, A
30	241	2.6	161	16 US-10-767-701-32672	Sequence 32672, A
31	202	2.2	585	15 US-10-424-599-285334	Sequence 285334,
32	188.5	2.0	2041	16 US-10-408-765A-1151	Sequence 1151, App
33	184	2.0	1979	14 US-10-205-823-419	Sequence 419, App
34	180.5	2.0	1413	15 US-10-369-493-6092	Sequence 6092, Ap
35	180	2.0	996	9 US-09-815-242-5251	Sequence 5251, Ap
36	180	2.0	1009	9 US-09-815-242-12141	Sequence 12141, A
37	179.5	1.9	2579	16 US-10-473-741-6	Sequence 6, Appli
38	179.5	1.9	2792	16 US-10-473-741-4	Sequence 4, Appli
39	179.5	1.9	2856	16 US-10-473-741-3	Sequence 3, Appli
40	178	1.9	1164	15 US-10-369-493-9770	Sequence 9770, Ap
41	177.5	1.9	2503	17 US-10-828-985A-11	Sequence 11, Appl
42	177.5	1.9	2543	17 US-10-828-985A-9	Sequence 9, Appli
43	177.5	1.9	2568	17 US-10-828-985A-7	Sequence 7, Appli
44	177.5	1.9	2816	15 US-10-240-145-145	Sequence 145, App
45	177	1.9	1151	16 US-10-408-765A-994	Sequence 994, App

#### ALIGNMENTS

##### RESULT 1

US-10-758-672A-2  
; Sequence 2, Application US/10758672A  
; Publication No. US20040185037A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966B  
; CURRENT APPLICATION NUMBER: US/10/758,672A  
; PRIOR FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 2  
; LENGTH: 1749  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-758-672A-2

Query Match 100.0%; Score 9224; DB 16; Length 1749;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAD	EAGGTERMEISAEPLQPTQRLASWDDQVFTAFHLHLAQLVPEIYFAEMDPDL	60
DB	1	MAD	EAGGTERMEISAEPLQPTQRLASWDDQVFTAFHLHLAQLVPEIYFAEMDPDL	60
QY	61	KOE	SVQMSITPLEWYLFGEEDPDLCKLKHGSAFQLCGRVFKSGTTCSCDCAIDPT	120
DB	61	KOE	SVQMSITPLEWYLFGEEDPDLCKLKHGSAFQLCGRVFKSGTTCSCDCAIDPT	120

121 QY CVLCMDCFODSVHKNHRYRMHTSTGGPCDCGDTAWKTGPFVNHGPRAGTIIKENSRC 180  
121 Db CVLCMDCFODSVHKNHRYRMHTSTGGPCDCGDTAWKTGPFVNHGPRAGTIIKENSRC 180  
181 QY PLNEEVIVQARKIPFSVIVYVEMTIIWEBEKELPPELOQIREKNERYCYVLFNDEHSHYDH 240  
181 Db PLNEEVIVQARKIPFSVIVYVEMTIIWEBEKELPPELOQIREKNERYCYVLFNDEHSHYDH 240  
241 QY VIYSQALDCELAELAHQHTTADKEGRRAVAGAYAAQOEAKEDIKSHSENVSHPLHV 300  
241 Db VIYSQALDCELAELAHQHTTADKEGRRAVAGAYAAQOEAKEDIKSHSENVSHPLHV 300  
301 QY EVLHSEIMAHQFALRLGSMWNKIMSYSDRFQIFCOACLRPEPDSNPCLLSRLMLMDA 360  
301 Db EVLHSEIMAHQFALRLGSMWNKIMSYSDRFQIFCOACLRPEPDSNPCLLSRLMLMDA 360  
361 QY KLYKGARKILHILIFSSPFMEYKCLFAMEFVKYKQLOKEYISDDHDSISITALSQV 420  
361 Db KLYKGARKILHILIFSSPFMEYKCLFAMEFVKYKQLOKEYISDDHDSISITALSQV 420  
421 QY MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFPGYSQDKLGRYAVICDLK 480  
421 Db MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFPGYSQDKLGRYAVICDLK 480  
481 QY YLILSKPTIWTBRLRMQFLEGRSPFLKILTCMOGMEIIRROYGOHIEVDPDWEAAIAIQM 540  
481 Db YLILSKPTIWTBRLRMQFLEGRSPFLKILTCMOGMEIIRROYGOHIEVDPDWEAAIAIQM 540  
541 QY QLKNIILMFQEWACDEBELLVAYKECHKAVMRCSTSFISSTKVVSQSGHSLKTSYRV 600  
541 Db QLKNIILMFQEWACDEBELLVAYKECHKAVMRCSTSFISSTKVVSQSGHSLKTSYRV 600  
601 QY SEDLSVILPLSRTLAGLHVRLSRGAVSRHLHFEVSPFQVEVLVYPLRCLVLVAQV 660  
601 Db SEDLSVILPLSRTLAGLHVRLSRGAVSRHLHFEVSPFQVEVLVYPLRCLVLVAQV 660  
661 QY AEMWRNGLSLSQVFPYQDVKCRBEMDKIIMLQIGASLMDPNKFLLLVLRVELABA 720  
661 Db AEMWRNGLSLSQVFPYQDVKCRBEMDKIIMLQIGASLMDPNKFLLLVLRVELABA 720  
721 QY FNKTISTKQDQLIKQNTLIEBMLQVLIYIVGERVYVPGVNTVGBEVTMREIHLICIEP 780  
721 Db FNKTISTKQDQLIKQNTLIEBMLQVLIYIVGERVYVPGVNTVGBEVTMREIHLICIEP 780  
781 QY MPHSAIAKLPENNETGLENNVINKVATPKPGVSGHGVYELKDBSLKDFNMFFHYSK 840  
781 Db MPHSAIAKLPENNETGLENNVINKVATPKPGVSGHGVYELKDBSLKDFNMFFHYSK 840  
841 QY TOHSAEHMOKRRKOENKDEALPPPPPPPCFAPSKVINLNCDDIMVILRTVPERAD 900  
841 Db TOHSAEHMOKRRKOENKDEALPPPPPPPCFAPSKVINLNCDDIMVILRTVPERAD 900  
901 QY TDSNLWTEGMLQWAFHILALGLLEKQOQKAPBEEVTPDFYHKASRLGSSAMNIOMLLE 960  
901 Db TDSNLWTEGMLQWAFHILALGLLEKQOQKAPBEEVTPDFYHKASRLGSSAMNIOMLLE 960  
961 QY KLKGPQLSGQKDMITWILQMFDTVRLREKSLIVATTSGSESKNDBITDKKKAERK 1020  
961 Db KLKGPQLSGQKDMITWILQMFDTVRLREKSLIVATTSGSESKNDBITDKKKAERK 1020  
1021 QY RXABAARLHROKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEBESTPAVSDYSRI 1080  
1021 Db RXABAARLHROKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEBESTPAVSDYSRI 1080  
1081 QY ALGPKRGPSVTEKEVLTCILQBEQEVKLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140  
1081 Db ALGPKRGPSVTEKEVLTCILQBEQEVKLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140  
1141 QY LDPLFMDPDLAYTGTYSCHVNHAVCWQKYFEAVOLSSQORIHVDLFDLESGEVILCPLC 1200  
1141 Db LDPLFMDPDLAYTGTYSCHVNHAVCWQKYFEAVOLSSQORIHVDLFDLESGEVILCPLC 1200  
1201 QY KSLCNTVPIPILOPQKINSENADALQLLTLARWIQTVLARISGVNIRHAKGENPIPIF 1260

1201 Db KSLCNTVPIPILOPQKINSENADALQLLTLARWIQTVLARISGVNIRHAKGENPIPIF 1260  
1261 QY FNQMGDSTLSPHSLSPGVESSIKYNSIKEMVILFATTIYRIGLVKPPDERDRPVML 1320  
1261 Db FNQMGDSTLSPHSLSPGVESSIKYNSIKEMVILFATTIYRIGLVKPPDERDRPVML 1320  
1321 QY TWSCAFTIOAIENLGDGKPLFCALONROHNGIKALMOFAVAQRITCPOVLIOKHLVR 1380  
1321 Db TWSCAFTIOAIENLGDGKPLFCALONROHNGIKALMOFAVAQRITCPOVLIOKHLVR 1380  
1381 QY LLSVVLPNIKSEDTPCLLSIDLFHVLVGAFLAPPSLYWDDPVDLPQSSVSSSYNHLYLHF 1440  
1381 Db LLSVVLPNIKSEDTPCLLSIDLFHVLVGAFLAPPSLYWDDPVDLPQSSVSSSYNHLYLHF 1440  
1441 QY LITMAHMLQIILLTVDTGLPLAQVQEDSBEAHSASFPABISQYTSGSIGCDIPGYLWVS 1500  
1441 Db LITMAHMLQIILLTVDTGLPLAQVQEDSBEAHSASFPABISQYTSGSIGCDIPGYLWVS 1500  
1501 QY LKNGITPVLRCALPFHYLLGVTPPELHTMSAEGEYSALCSYLSLPTNLLFOEYWD 1560  
1501 Db LKNGITPVLRCALPFHYLLGVTPPELHTMSAEGEYSALCSYLSLPTNLLFOEYWD 1560  
1561 QY VRPLLQRCADPALINCLKQKNTVVRYPKKNLSIELPDDYSCILNQASHFRCPSADDE 1620  
1561 Db VRPLLQRCADPALINCLKQKNTVVRYPKKNLSIELPDDYSCILNQASHFRCPSADDE 1620  
1621 QY RKHPVLCLFCGAILCSQNICQEIQVNGEVSACIFHALHCGAGVCIPLKIRECRVVLVEG 1680  
1621 Db RKHPVLCLFCGAILCSQNICQEIQVNGEVSACIFHALHCGAGVCIPLKIRECRVVLVEG 1680  
1681 QY KARGCAYPAPYLDEYGETDGLKGNPLHLSRERYKRLHVMQOHCIIETARSQETNOM 1740  
1681 Db KARGCAYPAPYLDEYGETDGLKGNPLHLSRERYKRLHVMQOHCIIETARSQETNOM 1740  
1741 QY LFGFNWQLL 1749  
1741 Db LFGFNWQLL 1749

RESULT 2

US-10-357-819-10  
; Sequence 10, Application US/10357819  
; Publication No. US20040259774A1  
; GENERAL INFORMATION:  
; APPLICANT: Alvarez, Enrique  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Gerlach, Valerie L.  
; APPLICANT: Gorman, Linda  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Li, Li  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Zhong, Mei  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-538A  
; CURRENT APPLICATION NUMBER: US/10/357,819  
; CURRENT FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: 09/520,781  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 09/584,411  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 09/783,436  
; PRIOR FILING DATE: 2001-02-14

```
; PRIOR APPLICATION NUMBER: 10/085,198
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/353,301
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/355,099
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/356,424
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/358,239
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,608
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/359,367
; PRIOR FILING DATE: 2002-02-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 10
; LENGTH: 1749
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-357-819-10

Query Match      100.0%; Score 9224; DB 16; Length 1749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADEBAGGTERMEISAELEPQTPORLASWMDQVDFYTAFLHLHLAQLVPEIYPAEMDPDLE 60
DB 1 MADEBAGGTERMEISAELEPQTPORLASWMDQVDFYTAFLHLHLAQLVPEIYPAEMDPDLE 60

QY 61 KQESVQMSIPTPLEWYLFGEPPDICLEKLKHSAPQLCGRVFKSGETTYSCRDCAIDPT 120
DB 61 KQESVQMSIPTPLEWYLFGEPPDICLEKLKHSAPQLCGRVFKSGETTYSCRDCAIDPT 120

QY 121 CVLQMDCFDSVHKNHRYKMHSTSGGFCDCGDTAEWKTPPCVNHPEGRAGTIKENSRC 180
DB 121 CVLQMDCFDSVHKNHRYKMHSTSGGFCDCGDTAEWKTPPCVNHPEGRAGTIKENSRC 180

QY 181 PLNEEVIVQARKIPPSVIKYVEMTWEEBKEPLPPELQIREKNERYCYLVFNDEHSHSYD 240
DB 181 PLNEEVIVQARKIPPSVIKYVEMTWEEBKEPLPPELQIREKNERYCYLVFNDEHSHSYD 240

QY 241 VIYSLQALDCELAEAQLHTTAIDKEGRAVAKAGAYAAQAEKEDIKSHSENVSHPLHV 300
DB 241 VIYSLQALDCELAEAQLHTTAIDKEGRAVAKAGAYAAQAEKEDIKSHSENVSHPLHV 300

QY 301 EVLHSEIMAHOKPALGLGSMNKNISYSDPRQIFCOACLRPEPDSNPCLISRLMLWDA 360
DB 301 EVLHSEIMAHOKPALGLGSMNKNISYSDPRQIFCOACLRPEPDSNPCLISRLMLWDA 360

QY 361 KLYKGARKILHELIFSSPFMEYKCLFAMEFVKYKQLOKEVYISDDHDSISITALSVO 420
DB 361 KLYKGARKILHELIFSSPFMEYKCLFAMEFVKYKQLOKEVYISDDHDSISITALSVO 420

QY 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFNGYQSKLGRVYVICDLK 480
DB 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFNGYQSKLGRVYVICDLK 480

QY 481 YILISKPTTWELRMQFLGFRSFLKILTCQGHIEIRROVGOHIEVDPDWEAAIAIQM 540
DB 481 YILISKPTTWELRMQFLGFRSFLKILTCQGHIEIRROVGOHIEVDPDWEAAIAIQM 540

QY 541 QLNKILLMFEQWACADELLLVAYKECHAVNRCSTSPISSTKVVSQSGHSLKTSYRV 600
DB 541 QLNKILLMFEQWACADELLLVAYKECHAVNRCSTSPISSTKVVSQSGHSLKTSYRV 600

QY 601 SEDLSVSIHLPLSRTLAGHLVRLSGLGAVSRLEHFEVSFQVEVLVYPLRCLVLVAQV 660
DB 601 SEDLSVSIHLPLSRTLAGHLVRLSGLGAVSRLEHFEVSFQVEVLVYPLRCLVLVAQV 660

QY 661 AEMWRNGLSLISQVFFYQDVVKREBMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 720
DB 661 AEMWRNGLSLISQVFFYQDVVKREBMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 720
```

```
DB 661 AEMWRNGLSLISQVFFYQDVVKREBMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 720
QY 721 FNKTTSTKDQDLIKQNTLIEBMLQVLIYVGERVYVPGVGNVTKBEVTWREIHLLCIEP 780
DB 721 FNKTTSTKDQDLIKQNTLIEBMLQVLIYVGERVYVPGVGNVTKBEVTWREIHLLCIEP 780
QY 781 MPHSIAKNLPENNENETGLENNVINKVATFKKPGVSGHGVVGLKDBSLKDFNMYFYHYSK 840
DB 781 MPHSIAKNLPENNENETGLENNVINKVATFKKPGVSGHGVVGLKDBSLKDFNMYFYHYSK 840
QY 841 TOHSAKAEKQKRRQENKDEALPPPPPEFCFAPFSKVINLNLNCDIMMYILRTVPERAID 900
DB 841 TOHSAKAEKQKRRQENKDEALPPPPPEFCFAPFSKVINLNLNCDIMMYILRTVPERAID 900
QY 901 TDSNLWTEGMLQMAFHIALGALLEBEKQLOKAPBESEVTFDFYHKASRLGSSAMNQLME 960
DB 901 TDSNLWTEGMLQMAFHIALGALLEBEKQLOKAPBESEVTFDFYHKASRLGSSAMNQLME 960
QY 961 KLKGIPOLEGOKDMITWLOMPDVTVKRLREKSLCLVATTSSESITKNDIETHDKBAERK 1020
DB 961 KLKGIPOLEGOKDMITWLOMPDVTVKRLREKSLCLVATTSSESITKNDIETHDKBAERK 1020
QY 1021 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEBESTPAVDYSRI 1080
DB 1021 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEBESTPAVDYSRI 1080
QY 1081 ALGPKRGPSTVEKEVLTCTILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
DB 1081 ALGPKRGPSTVEKEVLTCTILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
QY 1141 LDPLFMDPDLAYGTTGSCGHVHMAVCWQKYPEAVOLSSQQRHIVDLFDESGEVLCPLC 1200
DB 1141 LDPLFMDPDLAYGTTGSCGHVHMAVCWQKYPEAVOLSSQQRHIVDLFDESGEVLCPLC 1200
QY 1201 KSLCNTVPIPIPLQPKINSENADALAOULLTLARWITQTLARISGVNIRHAKGENPIPIF 1260
DB 1201 KSLCNTVPIPIPLQPKINSENADALAOULLTLARWITQTLARISGVNIRHAKGENPIPIF 1260
QY 1261 FNQMGDSTLBFHSTLSFGVSESSIKYSIKEMVILPATIYIRIGLKVPPDPRVPM 1320
DB 1261 FNQMGDSTLBFHSTLSFGVSESSIKYSIKEMVILPATIYIRIGLKVPPDPRVPM 1320
QY 1321 TWSTCAFTIOAENILGDEGKPLFCALONRQHNGIKALMQFAVQORITCPOVLIOKHLVR 1380
DB 1321 TWSTCAFTIOAENILGDEGKPLFCALONRQHNGIKALMQFAVQORITCPOVLIOKHLVR 1380
QY 1381 LLSVVLVPLNKSDETPCLLSIDLPHVLVGAFLAPPSLYWDDPVDLOPSSVSSSYNHLXLP 1440
DB 1381 LLSVVLVPLNKSDETPCLLSIDLPHVLVGAFLAPPSLYWDDPVDLOPSSVSSSYNHLXLP 1440
QY 1441 LITMAHMLQILLTVDTGLPLAQVQDSSEAHSAFFAEISQYTSGSIGCDIPGYLWVS 1500
DB 1441 LITMAHMLQILLTVDTGLPLAQVQDSSEAHSAFFAEISQYTSGSIGCDIPGYLWVS 1500
QY 1501 LKNGITPYLRCAALPFHYLLGVTTPPELHTNSAEGEYSALCSYLSLPTNLFLFOEYWD 1560
DB 1501 LKNGITPYLRCAALPFHYLLGVTTPPELHTNSAEGEYSALCSYLSLPTNLFLFOEYWD 1560
QY 1561 VRPLQRCADPALINCLKQNTVYRPRKNSLIELPDDYSCLANQASHPRCPASADDE 1620
DB 1561 VRPLQRCADPALINCLKQNTVYRPRKNSLIELPDDYSCLANQASHPRCPASADDE 1620
QY 1621 RKHPVCLFCGAILCSQNICCOEIVNGEVEGACIFHALHCGAGVCIPLKIRECRVVLVEG 1680
DB 1621 RKHPVCLFCGAILCSQNICCOEIVNGEVEGACIFHALHCGAGVCIPLKIRECRVVLVEG 1680
QY 1681 KARGCAYPAPYLDVEGETDPLKRGKPNLHLSRERYRKLHLVWQHCIIIEIARSETNQ 1740
DB 1681 KARGCAYPAPYLDVEGETDPLKRGKPNLHLSRERYRKLHLVWQHCIIIEIARSETNQ 1740
QY 1741 LFGFNWQLL 1749
DB 1741 LFGFNWQLL 1749
```

RESULT 3									
US-10-758-636A-2									
; Sequence 2, Application US/10758636A									
; Publication No. US20050089876A1									
; GENERAL INFORMATION:									
; APPLICANT: Han, et al.									
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY									
; FILE REFERENCE: 01017/35966C									
; CURRENT APPLICATION NUMBER: US/10/758,636A									
; CURRENT FILING DATE: 2004-01-15									
; PRIOR APPLICATION NUMBER: US 09/724,126									
; PRIOR FILING DATE: 2000-11-28									
; PRIOR APPLICATION NUMBER: US 60/187,911									
; PRIOR FILING DATE: 2000-03-08									
; NUMBER OF SEQ ID NOS: 29									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 2									
; LENGTH: 1749									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-758-636A-2									
Query Match 100.0%; Score 9224; DB 17; Length 1749;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MADEAGGTERMEISAEFLPQTQRLASWDDQVDFYTAFLHLAQLVPEIYPAEMDPDLE	60						
DB	1	MADEAGGTERMEISAEFLPQTQRLASWDDQVDFYTAFLHLAQLVPEIYPAEMDPDLE	60						
QY	61	KQESVQMSIPTLEWYLFGEPPDICLEKXUSGAPQLCGRVFKSGETTYSRCDCAIDPT	120						
DB	61	KQESVQMSIPTLEWYLFGEPPDICLEKXUSGAPQLCGRVFKSGETTYSRCDCAIDPT	120						
QY	121	CVLQMDCFQDSVHKHRYKWHSTGGGCDGTEAWKTGPPCVNHEPGRAGTIKENSRC	180						
DB	121	CVLQMDCFQDSVHKHRYKWHSTGGGCDGTEAWKTGPPCVNHEPGRAGTIKENSRC	180						
QY	181	PLNEEVIQARKIFPSVIVYVEMTIWEEBKEPLPELQIREKNERYCYLVFNDHHSYDH	240						
DB	181	PLNEEVIQARKIFPSVIVYVEMTIWEEBKEPLPELQIREKNERYCYLVFNDHHSYDH	240						
QY	241	VIYSQRALDCELAHQHTTALDKGRRAVKAGYAAQCBKEDIKSHSENVSOHPLHV	300						
DB	241	VIYSQRALDCELAHQHTTALDKGRRAVKAGYAAQCBKEDIKSHSENVSOHPLHV	300						
QY	301	EVLHSEIMAHQFALRLGSMWNKIMSYSDFRQIFCOACLRBEPDSENPCLISRLMLWA	360						
DB	301	EVLHSEIMAHQFALRLGSMWNKIMSYSDFRQIFCOACLRBEPDSENPCLISRLMLWA	360						
QY	361	KLYKGARKILHELIFSSFMEYKGLFAMEFVKYKQLQKEYISDDHDSISITALSQV	420						
DB	361	KLYKGARKILHELIFSSFMEYKGLFAMEFVKYKQLQKEYISDDHDSISITALSQV	420						
QY	421	MFTVPTLARHLIEBQNVISVITETLLVPLPYLDNRNKNFQGYSDKLGRVYVICDLK	480						
DB	421	MFTVPTLARHLIEBQNVISVITETLLVPLPYLDNRNKNFQGYSDKLGRVYVICDLK	480						
QY	481	YILISKPTIWTLELRMQFLGFRSFLKILTCMQGMEIRROVGOHLEVPDMEAAIAIQM	540						
DB	481	YILISKPTIWTLELRMQFLGFRSFLKILTCMQGMEIRROVGOHLEVPDMEAAIAIQM	540						
QY	541	QLKNILLMFQEWACDDEELLVAYKECHKAVMRCSTSFISSSKTVVQSCGHSLETKSYRV	600						
DB	541	QLKNILLMFQEWACDDEELLVAYKECHKAVMRCSTSFISSSKTVVQSCGHSLETKSYRV	600						
QY	601	SEDLVSIHPLSRTLAGLHVRLSRGAVSRLHEFVSFDPQVEVLVEYPLRCLVLVAQVY	660						
DB	601	SEDLVSIHPLSRTLAGLHVRLSRGAVSRLHEFVSFDPQVEVLVEYPLRCLVLVAQVY	660						
QY	661	AEWWRNGLSLISQVFFYQDVKCREEMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA	720						



Db 1741 LFGFNWQLL 1749

RESULT 4  
US-10-287-218-1  
; Sequence 1, Application US/10287218  
; Publication No. US20030198975A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.  
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.  
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.  
; APPLICANT: BUFORD, Neil; DING, Li  
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.  
; APPLICANT: GANDHI, Ameena R.; GIBTZEN, Kimberly J.  
; APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.  
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.  
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.  
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi  
; APPLICANT: REDDY, Roopa; SANJANWALA, Madhu, M.  
; APPLICANT: TANG, Y. Tom; WALIA, Narinder K.  
; APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.  
; APPLICANT: XU, Yuming; YANG, Junning  
; APPLICANT: YAO, Monique G.; YUE, Henry  
; APPLICANT: ZEBARJAN, Yeganeh  
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH  
; FILE REFERENCE: PI-0417 USA  
; CURRENT APPLICATION NUMBER: US/10/287,218  
; PRIORITY FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: PCT/US02/11152  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: US 60/349,705  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: US 60/295,263  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/295,340  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/293,727  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 60/291,846  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/291,662  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: US 60/287,228  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/286,820  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/283,294  
; PRIOR FILING DATE: 2001-04-11  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 1738  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030198975A1 1351608CD1  
US-10-287-218-1

Query Match 99.4%; Score 9168; DB 14; Length 1738;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MEISAEIPOTPORLASWMDQVDFTYAFHLHLAQLVPEIYFAEMDPDLKQESVQMSTP 71  
DB 1 MEISAEIPOTPORLASWMDQVDFTYAFHLHLAQLVPEIYFAEMDPDLKQESVQMSTP 60  
QY 72 TPLEWYLFGEDEPDICLEKUKHSGAFQLCGRVFKSGRTTYSRCDCAIDPTCVLCMDCFQDS 131  
DB 61 TPLEWYLFGEDEPDICLEKUKHSGAFQLCGRVFKSGRTTYSRCDCAIDPTCVLCMDCFQDS 120

QY 132 VHKHRYXKHTSTGCGFCDCGTBAWKTGPPCVNHEPGRAGTIKENSRCPLNEEVIQAR 191  
DB 121 VHKHRYXKHTSTGCGFCDCGTBAWKTGPPCVNHEPGRAGTIKENSRCPLNEEVIQAR 180  
QY 192 KIFPSVIKYVEMTWEKELPELQIREKNERYCVLFNDEHSHSDHVIYSLOALDQ 251  
DB 181 KIFPSVIKYVEMTWEKELPELQIREKNERYCVLFNDEHSHSDHVIYSLOALDQ 240  
QY 252 ELAEALHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSHPLHVEVLHSEIMAHQ 311  
DB 241 ELAEALHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSHPLHVEVLHSEIMAHQ 300  
QY 312 KPALRLGSMWNKIMSYSDFRQIFQACLRPEPDSNPCLISRLMLWDALYKGAARKILH 371  
DB 301 KPALRLGSMWNKIMSYSDFRQIFQACLRPEPDSNPCLISRLMLWDALYKGAARKILH 360  
QY 372 ELIFSSPFMEYKCLFAMEFVKYKQKQKYSIDHDSRISITALSVMQPTVPTLARHL 431  
DB 361 ELIFSSPFMEYKCLFAMEFVKYKQKQKYSIDHDSRISITALSVMQPTVPTLARHL 420  
QY 432 IBEQNVISVITETLLEVLPEYLDNRNKNFQGSQDKLGRVYAVICDLKYILISKPTWT 491  
DB 421 IBEQNVISVITETLLEVLPEYLDNRNKNFQGSQDKLGRVYAVICDLKYILISKPTWT 480  
QY 492 ERLRQPLEGFRSLKILTCQMEIRROVQCHIEVDPDWEAAIAIQMLKNVILMFOE 551  
DB 481 ERLRQPLEGFRSLKILTCQMEIRROVQCHIEVDPDWEAAIAIQMLKNVILMFOE 540  
QY 552 WCACDEBELLVAYKECHKAVMRCSTFISSTKVTVQSCGHSLETYSRVSEDLVSHLPL 611  
DB 541 WCACDEBELLVAYKECHKAVMRCSTFISSTKVTVQSCGHSLETYSRVSEDLVSHLPL 600  
QY 612 SRTLAGLHVRLSRLGAVSRLEHFEVDFQVEVLVEYPLRCLVLVAQVVAEMWRENGSL 671  
DB 601 SRTLAGLHVRLSRLGAVSRLEHFEVDFQVEVLVEYPLRCLVLVAQVVAEMWRENGSL 660  
QY 672 ISQVFFYQDVKREEMDYDKIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTISTKDD 731  
DB 661 ISQVFFYQDVKREEMDYDKIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTISTKDD 720  
QY 732 LIKOYNTLIEBMLQVLIYVGERVYVGVNVTKEBVTWREIHLACIEMPHSAKONLP 791  
DB 721 LIKOYNTLIEBMLQVLIYVGERVYVGVNVTKEBVTWREIHLACIEMPHSAKONLP 780  
QY 792 ENENNETGLENVINKVATFKPGVSGHGYELKDESLKDFNMYFYHYSKTKHSAEHMQ 851  
DB 781 ENENNETGLENVINKVATFKPGVSGHGYELKDESLKDFNMYFYHYSKTKHSAEHMQ 840  
QY 852 KRRKQENKDEALPPPPPPPCFAPSKVINLNCIDIMMVLRTVFPFRAIDTDSNLTGML 911  
DB 841 KRRKQENKDEALPPPPPPPCFAPSKVINLNCIDIMMVLRTVFPFRAIDTDSNLTGML 900  
QY 912 QMAFHILALGLLEKQLOKAPBEVTFDFVHKASRLGSSAMNIOMLKLGIPQESQ 971  
DB 901 QMAFHILALGLLEKQLOKAPBEVTFDFVHKASRLGSSAMNIOMLKLGIPQESQ 960  
QY 972 KOMITWILQMFDTVKRLKESCLIVATTSGSEIKNDEITHDKKAEKRAEAAARLHQ 1031  
DB 961 KOMITWILQMFDTVKRLKESCLIVATTSGSEIKNDEITHDKKAEKRAEAAARLHQ 1020  
QY 1032 KIMAQMSALQKNPIETHKLMYDNTSEMPGKDSIMEESTPAVSYSRIALGPGRGPSVT 1091  
DB 1021 KIMAQMSALQKNPIETHKLMYDNTSEMPGKDSIMEESTPAVSYSRIALGPGRGPSVT 1080  
QY 1092 EKEVLTCILCOBQEVKLENNAMVLSACVQKSTALTOHRGKPIELSGALPDLMDPLA 1151  
DB 1081 EKEVLTCILCOBQEVKLENNAMVLSACVQKSTALTOHRGKPIELSGALPDLMDPLA 1140  
QY 1152 YGTGTSGCHVHAWCQKYPEAVQLSQQRHVDLFDLESGEVLCPLCKSLCNTVPII 1211  
DB 1141 YGTGTSGCHVHAWCQKYPEAVQLSQQRHVDLFDLESGEVLCPLCKSLCNTVPII 1200  
QY 1212 PLQPOKINSNADALALQLTLARWITQTVLARI SGYNIRHAKGENPIPIFFNQMGDSTLE 1271

Db 1201 PLOPQKINSENADALQALLTLARWLTQVLAISGYNIRHAKGENPIPIFFNQGMDSTLE 1260  
QY 1272 FHSILSPGVSESIKYSNSIKEMVILFATTIYRIGLKVPDPDRPRVPMLTWSTCAFTQA 1331  
Db 1261 FHSILSPGVSESIKYSNSIKEMVILFATTIYRIGLKVPDPDRPRVPMLTWSTCAFTQA 1320  
QY 1332 IENLLGDEKPLFGALONRQNGKALAMQFAVAQRITCPQVLIQKHLVRLSVLPNTKS 1391  
Db 1321 IENLLGDEKPLFGALONRQNGKALAMQFAVAQRITCPQVLIQKHLVRLSVLPNTKS 1380  
QY 1392 EPTPCLLSIDLFLVAVLAPPSLYWDDPVDLPQSSVSSYNHLYLPHLITMAHMLQIL 1451  
Db 1381 EPTPCLLSIDLFLVAVLAPPSLYWDDPVDLPQSSVSSYNHLYLPHLITMAHMLQIL 1440  
QY 1452 LTVDTGLPLAQVQDSSEAHSSAFPAETISQVTSIGSDIPGWYLMVSLKNGITPYLR 1511  
Db 1441 LTVDTGLPLAQVQDSSEAHSSAFPAETISQVTSIGSDIPGWYLMVSLKNGITPYLR 1500  
QY 1512 ALPFLHYLGVTPPELHNTSABGEYSALCSYLSPTNLFLPQBYWDTVRLLQWCAD 1571  
Db 1501 ALPFLHYLGVTPPELHNTSABGEYSALCSYLSPTNLFLPQBYWDTVRLLQWCAD 1560  
QY 1572 PALLNCLQKQNTVRYPRKRNLSIELPDDYSCLLQASHFRCPRSADDERKHPVLCFCG 1631  
Db 1561 PALLNCLQKQNTVRYPRKRNLSIELPDDYSCLLQASHFRCPRSADDERKHPVLCFCG 1620  
QY 1632 AILCSQNTCCQIRVNGEEVGACIFHALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPY 1691  
Db 1621 AILCSQNTCCQIRVNGEEVGACIFHALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPY 1680  
QY 1692 LDEYGETDPLGRGNPLHLRSRYRKLHLVWQOHCIIIEIARSQETNOMLFGFNWQLL 1749  
Db 1681 LDEYGETDPLGRGNPLHLRSRYRKLHLVWQOHCIIIEIARSQETNOMLFGFNWQLL 1738

## RESULT 5

US-10-474-291-1  
; Sequence 1, Application US/10474291  
; Publication No. US20040132043A1  
; GENERAL INFORMATION:  
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.  
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.  
; APPLICANT: BECHA, Shanya D.; BOWSKY, Mark L.  
; APPLICANT: BURFORD, Neil; DING, Li  
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.  
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.  
; APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.  
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.  
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.  
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi  
; APPLICANT: REDDY, Roopa M.; SANJANWALA, Madhusudan M.  
; APPLICANT: TANG, Y. Tom; CHAWLA, Naxinder K.  
; APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.  
; APPLICANT: XU, Yuming; YANG, Junming  
; APPLICANT: YAO, Monique G.; YUE, Henry  
; APPLICANT: ZEBARJADIAN, Vaganeh  
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH  
; FILE REFERENCE: PI-0417 USN  
; CURRENT APPLICATION NUMBER: US/10/474,291  
; CURRENT FILING DATE: 2003-10-06  
; PRIOR APPLICATION NUMBER: PCT/US02/11152  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: US 60/349,705  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: US 60/295,263  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/295,340  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US 60/293,727  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 60/291,846  
; PRIOR FILING DATE: 2001-05-18

; PRIOR APPLICATION NUMBER: US 60/291,662  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: US 60/287,228  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/286,820  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/283,294  
; PRIOR FILING DATE: 2001-04-11  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 1738  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 1351608CD1  
US-10-474-291-1

Query Match 99.4%; Score 9168; DB 16; Length 1738;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 MEISAEPLPOTPORLASWMDQVDFYTAFLHHLAQLVPEIYFAEMDPDLKQEEVSQMSIF 71  
Db 1 MEISAEPLPOTPORLASWMDQVDFYTAFLHHLAQLVPEIYFAEMDPDLKQEEVSQMSIF 60  
QY 72 TPLEWYLFGEPPDICLEKLKHSQAPQLCGRVFKSGETTYSRCDCAIDPTCVLCMDCFQDS 131  
Db 61 TPLEWYLFGEPPDICLEKLKHSQAPQLCGRVFKSGETTYSRCDCAIDPTCVLCMDCFQDS 120  
QY 132 VHKHRYKMHSTSGGFCDCGTEAWKTGPPCVNHPGPRAGTIKENSRCPLNEEVIQAR 191  
Db 121 VHKHRYKMHSTSGGFCDCGTEAWKTGPPCVNHPGPRAGTIKENSRCPLNEEVIQAR 180  
QY 192 KIPPSVIKVVVEMTIWEEBKEPLPELOQIREKNERYCYLVFNDEHHSYDHVYISLQALDC 251  
Db 191 KIPPSVIKVVVEMTIWEEBKEPLPELOQIREKNERYCYLVFNDEHHSYDHVYISLQALDC 240  
QY 252 ELAERLAQLHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSOHPLHVEVLHSGEIMAHQ 311  
Db 241 ELAERLAQLHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSOHPLHVEVLHSGEIMAHQ 300  
QY 312 KPALRLGSMWNKIMSYSDFRQIFCOACLRREPDSNPCLISRLMLWDAKLYKGARKILH 371  
Db 301 KPALRLGSMWNKIMSYSDFRQIFCOACLRREPDSNPCLISRLMLWDAKLYKGARKILH 360  
QY 372 ELIFSSPFMEYKXKLFAMEFVKYKQLOKEYISDDHRSISITALSVMQMTVPTLARHL 431  
Db 361 ELIFSSPFMEYKXKLFAMEFVKYKQLOKEYISDDHRSISITALSVMQMTVPTLARHL 420  
QY 432 IEEQNVI SVITETTLLEVLPEYLDNRNKNFNGYSGDKLGRVYAVICDLKYILISKPTIWT 491  
Db 421 IEEQNVI SVITETTLLEVLPEYLDNRNKNFNGYSGDKLGRVYAVICDLKYILISKPTIWT 480  
QY 492 ERLRMQFLGFRSFLKILTCMOGHEEIRROVQOHEIYDPPDWEAAIAIQMLKNILLMPQE 551  
Db 481 ERLRMQFLGFRSFLKILTCMOGHEEIRROVQOHEIYDPPDWEAAIAIQMLKNILLMPQE 540  
QY 552 WCACDEEELLVAYKECHKAVMRCSTSFISSTKTVVQSCGHSLETYSKYSVSDLSIHLPL 611  
Db 541 WCACDEEELLVAYKECHKAVMRCSTSFISSTKTVVQSCGHSLETYSKYSVSDLSIHLPL 600  
QY 612 SRTLGLHVLRLSRLGAVSRLEHFVSFEDQVEVLVEYPLRCLVLVAQVVAEMWRNGLSL 671  
Db 601 SRTLGLHVLRLSRLGAVSRLEHFVSFEDQVEVLVEYPLRCLVLVAQVVAEMWRNGLSL 660  
QY 672 ISQVPYQDVVKCREEMVDKDIIMLQIGASLMDPNKFLLLVQLQRYELAEAFNKTIISTKOD 731  
Db 661 ISQVPYQDVVKCREEMVDKDIIMLQIGASLMDPNKFLLLVQLQRYELAEAFNKTIISTKOD 720  
QY 732 LIKQYNTLIEBMLQVLIYIGERYVPGVGNVTKBEVTWREIHLHCIBFMPHSAIAQNLP 791

Db 721 LIKQYNTLIEMLQVLIYVGERYVGVGNVTKERTMREIHLCLIEPMHSAIAKNLP 780  
QY 792 ENENNETGLENVINKVATFKKPGVSGHGVYELKDESXKDPFNMVYFYHYSKTQHSKAEHMOK 851  
Db 781 ENENNETGLENVINKVATFKKPGVSGHGVYELKDESXKDPFNMVYFYHYSKTQHSKAEHMOK 840  
QY 852 KRRKQENKDEALPPPPPPPCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLTGML 911  
Db 841 KRRKQENKDEALPPPPPPPCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLTGML 900  
QY 912 QNAFHILALGLLEKQKQLOKAPBEVTFDYFKASRLGSSAMNIOMLLEKLGIPQLEQ 971  
Db 901 QNAFHILALGLLEKQKQLOKAPBEVTFDYFKASRLGSSAMNIOMLLEKLGIPQLEQ 960  
QY 972 KOMITWILQWFTVTKRLREKSCLIIVATTSGSEISKNDEITHDKKAEKRKKAABARLHQ 1031  
Db 961 KOMITWILQWFTVTKRLREKSCLIIVATTSGSEISKNDEITHDKKAEKRKKAABARLHQ 1020  
QY 1032 KIMAQMSALQKQPIETHKLMYDNTSEMPGKEDSIMEESTPAVDYSRIALGPGRGPSVT 1091  
Db 1021 KIMAQMSALQKQPIETHKLMYDNTSEMPGKEDSIMEESTPAVDYSRIALGPGRGPSVT 1080  
QY 1092 EKEVLTCILCOBEOBKVINNAWLSACVOKSTALTQHRGKPIELSGEALDPLFMDPDLA 1151  
Db 1081 EKEVLTCILCOBEOBKVINNAWLSACVOKSTALTQHRGKPIELSGEALDPLFMDPDLA 1140  
QY 1152 YGTYTSCGHVHVAWCQKYFEAVQLSSQRIHVDLPDLESEGYLCPLCKSLCNTVPII 1211  
Db 1141 YGTYTSCGHVHVAWCQKYFEAVQLSSQRIHVDLPDLESEGYLCPLCKSLCNTVPII 1200  
QY 1212 PLOPQKINSENADALQTLTARWIOTVLARISGVNIRHAKENPIPIPFNGMGDSTLE 1271  
Db 1201 PLOPQKINSENADALQTLTARWIOTVLARISGVNIRHAKENPIPIPFNGMGDSTLE 1260  
QY 1272 FHSILSPGVESSTIKYNSIKEMVILPATIYRIGLKVPPDERDPRVPLMTWTCAFTIOA 1331  
Db 1261 FHSILSPGVESSTIKYNSIKEMVILPATIYRIGLKVPPDERDPRVPLMTWTCAFTIOA 1320  
QY 1332 IENLLGDEGKPLFGALQNRQHNGLKALMQFAVAQRITCPQVLTQKHLVRLLSVLPNKS 1391  
Db 1321 IENLLGDEGKPLFGALQNRQHNGLKALMQFAVAQRITCPQVLTQKHLVRLLSVLPNKS 1380  
QY 1392 EDTPCLLSIDLPHVLVGAVALPSPSYWDDPVDLPQSSVSSSYNNHLYLPHLITMAHMLQIL 1451  
Db 1381 EDTPCLLSIDLPHVLVGAVALPSPSYWDDPVDLPQSSVSSSYNNHLYLPHLITMAHMLQIL 1440  
QY 1452 LTVDTGLPLAQOEDSEEAHSAASPABISQYTSIGCDIPGWYLVSLKNGITPYLRC 1511  
Db 1441 LTVDTGLPLAQOEDSEEAHSAASPABISQYTSIGCDIPGWYLVSLKNGITPYLRC 1500  
QY 1512 AALFPHYLGVTPPEELHTNSABGEYSALCSYLSLPTNLPLFOEYWDVTRPLLQWCAD 1571  
Db 1501 AALFPHYLGVTPPEELHTNSABGEYSALCSYLSLPTNLPLFOEYWDVTRPLLQWCAD 1560  
QY 1572 PALLNCLKQKNTVVRYPKRNLSIELPDDYSCLLQAASHFRCPRSADDERKHPVLCFG 1631  
Db 1561 PALLNCLKQKNTVVRYPKRNLSIELPDDYSCLLQAASHFRCPRSADDERKHPVLCFG 1620  
QY 1632 AILCSQNTCCQIBVNGEEVGACIFALHCGAGVCIFLKIREDRVVLVEGKARGCAYPAPY 1691  
Db 1621 AILCSQNTCCQIBVNGEEVGACIFALHCGAGVCIFLKIREDRVVLVEGKARGCAYPAPY 1680  
QY 1692 LDEYGETDPLKRGKGNPLHLSRERYRKLHLVWOQHCHLIEIARSQETNOMLFGFNWOLL 1749  
Db 1681 LDEYGETDPLKRGKGNPLHLSRERYRKLHLVWOQHCHLIEIARSQETNOMLFGFNWOLL 1738

## RESULT 6

US-10-758-672A-19

; Sequence 19, Application US/10758672A

; Publication No. US20040185037A1

; GENERAL INFORMATION:

; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966B  
; CURRENT APPLICATION NUMBER: US/10/758,672A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 19  
; LENGTH: 1734  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-758-672A-19

Query Match 98.5%; Score 9087; DB 16; Length 1734;

Best Local Similarity 98.8%; Pred. No. 0;

Matches 1728; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MADEAGGTERMEISAEILPQTTPQRLASWDDQQVDYFATFLHHLAQLVPEIYFAEMDPDL 60  
Db 6 MADEAGGTERMEISAEILPQTTPQRLASWDDQQVDYFATFLHHLAQLVPEIYFAEMDPDL 65  
QY 61 KOEESVQMSIFTPLEWYLFGBDDPDI CLKELKHSGAPQLCGRVFKSGETTYSRCDCAIDPT 120  
Db 66 KOEESVQMSIFTPLEWYLFGBDDPDI CLKELKHSGAPQLCGRVFKSGETTYSRCDCAIDPT 125  
QY 121 CVLCHMDCPQDSVHKHNRKYMHTSTGGFCDCGDTBAWTKGPCVNHHEPGRAGTIKENSRC 180  
Db 126 CVLCHMDCPQDSVHKHNRKYMHTSTGGFCDCGDTBAWTKGPCVNHHEPGRAGTIKENSRC 185  
QY 181 PLNEBEVIVQARKIPFSVIKYVEMVTIWEKEKELPELQIREKNERYCYVLFNDEHSHYDH 240  
Db 186 PLNEBEVIVQARKIPFSVIKYVEMVTIWEKEKELPELQIREKNERYCYVLFNDEHSHYDH 245  
QY 241 VIYSLQALDCELAQAQLHTTAIDKEGRRAVKAGAYAAQCEAKEDIKSHSENVSOHPLHV 300  
Db 246 VIYSLQALDCELAQAQLHTTAIDKEGRRAVKAGAYAAQCEAKEDIKSHSENVSOHPLHV 305  
QY 301 EYVLSHMAHQFALRGCSWNNKINSYSDPRQIFCOACLRPEPSENPCILSRMLMWA 360  
Db 306 EYVLSHMAHQFALRGCSWNNKINSYSDPRQIFCOACLRPEPSENPCILSRMLMWA 365  
QY 361 KLYKGARKILHELIFSSPFMEYKCLPAMEFVKYKQLOKEYISDDHDSISITALSVO 420  
Db 366 KLYKGARKILHELIFSSPFMEYKCLPAMEFVKYKQLOKEYISDDHDSISITALSVO 425  
QY 421 MFTVPTLARHLIEQNVI SVITETLLEVLPEYLDNRNKNFQGSQDKLGRVYAVICDLK 480  
Db 426 MFTVPTLARHLIEQNVI SVITETLLEVLPEYLDNRNKNFQGSQDKLGRVYAVICDLK 485  
QY 481 YILISKPTIWTRELRMQPLEGFRSPFLKILTCMOGNEEIRROVQGHIEYVDPWEAAIAIQM 540  
Db 486 YILISKPTIWTRELRMQPLEGFRSPFLKILTCMOGNEEIRROVQGHIEYVDPWEAAIAIQM 545  
QY 541 QLNKILLMFQSCACDEELLVAVYKCHKAVNRCTSFISSTKVQSCGSHLETYSRV 600  
Db 546 QLNKILLMFQSCACDEELLVAVYKCHKAVNRCTSFISSTKVQSCGSHLETYSRV 605  
QY 601 SEDLVS IHLPLSRTLGLHVLRLSRLGAVSRHLHEFVSFEDFQVLEVLYEPLCLVLVAQVV 660  
Db 606 SEDLVS IHLPLSRTLGLHVLRLSRLGAVSRHLHEFVSFEDFQVLEVLYEPLCLVLVAQVV 665  
QY 661 AEMWRRNGLSLISQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELABA 720  
Db 666 AEMWRRNGLSLISQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELABA 725  
QY 721 FNKTI STKQDQLIKOYNTLIEMLQVLIYVGERYVGVGNVTKERTMREIHLCLIEP 780  
Db 726 FNKTI STKQDQLIKOYNTLIEMLQVLIYVGERYVGVGNVTKERTMREIHLCLIEP 785

781 MPHSALAKNLPENNETGLENVINKVATFKKPGVSGHGVYELKDESKDFMFFHYSK 840  
786 MPHSALAKNLPENNETGLENVINKVATFKKPGVSGHGVYELKDESKDFMFFHYSK 845  
841 TQHSKAEHQKRRKOENKDEALPPPPFPCFAPSKVINLNCDDIMYILRTVPERAID 900  
846 TQHSKAEHQKRRKOENKDEALPPPPFPCFAPSKVINLNCDDIMYILRTVPERAID 905  
901 TDSNLWTEGMLQMAFHILALGLEKQLOKAPREEVTFDFVHKASRLGSSAMNOMLLE 960  
906 TDSNLWTEGMLQMAFHILALGLEKQLOKAPREEVTFDFVHKASRLGSSAMNOMLLE 965  
961 KLKGTIPOLGQKDMITWILQMDPTVKRLREKSLVATTSSESINKDEITHDKKARK 1020  
966 KLKGTIPOLGQKDMITWILQMDPTVKRLREKSLVATTSSESINKDEITHDKKARK 1025  
1021 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEESTPAVSDYSRI 1080  
1026 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEESTPAVSDYSRI 1085  
1081 ALGPKRGPSVTEKEVLTCLCOBEOBKVLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140  
1086 ALGPKRGPSVTEKEVLTCLCOBEOBKVLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1145  
1141 LDPLFMDPLAYGTGTGSCGHVHVAWCQKYPEAVQLSSQOQRIHVDLPDLESGEYLCPIC 1200  
1146 LDPLFMDPLAYGTGTGSCGHVHVAWCQKYPEAVQLSSQOQRIHVDLPDLESGEYLCPIC 1205  
1201 KSLCNTVPIIIPLOPKINSENADALQALLTLARMIQTVLARISGYNIRHAKGENPIPIF 1260  
1206 KSLCNTVPIIIPLOPKINSENADALQALLTLARMIQTVLARISGYNIRHAKGENPIPIF 1265  
1261 FNQGMGDSLEPHSILSPGVESIKYSNISKEMVILFATTIYRIGLKVPDPDRDPRVPM 1320  
1266 FNQGMGDSLEPHSILSPGVESIKYSNISKEMVILFATTIYRIGLKVPDPDRDPRVPM 1325  
1321 TWSTCAFTIOALENLGDGKPLFGALQNRQNGKALMQPVAQRITCPQVLIQKHLVR 1380  
1326 TWSTCAFTIOALENLGDGKPLFGALQNRQNGKALMQPVAQRITCPQVLIQKHLVR 1385  
1381 LLSVLPNPKSBDTPCLLSIDLPHVLVGAVALPSPSYMDPDVLDLPSSVSSSYNHLYLPH 1440  
1386 LLSVLPNPKSBDTPCLLSIDLPHVLVGAVALPSPSYMDPDVLDLPSSVSSSYNHLYLPH 1445  
1441 LITMAHMLQILTVDTGLPLAQVEDSEBAHSASPABISQYTSGSTCDDIPGHWLWS 1500  
1446 LITMAHMLQILTVDTGLPLAQVEDSEBAHSASPABISQYTSGSTCDDIPGHWLWS 1505  
1501 LKNGITPYLRCAALPFFHYLLGVTPPELHTNAGBYSALCSYLSLPTNLPLLFQBYWDT 1560  
1506 LKNGITPYLRCAALPFFHYLLGVTPPELHTNAGBYSALCSYLSLPTNLPLLFQBYWDT 1565  
1561 VRPLLQRCADPALLNCLQKNTVVYPRKRNLSIELPDDYSCLLNQASHFRCPRSADDE 1620  
1566 VRPLLQRCADPALLNCLQKNTVVYPRKRNLSIELPDDYSCLLNQASHFRCPRSADDE 1625  
1621 RKHPVLCLFCGATLCSQNTCCQIBNVGBEVGACIFHNLHGAGVCIFPLKRECRVVLVEG 1680  
1626 RKHPVLCLFCGATLCSQNTCCQIBNVGBEVGACIFHNLHGAGVCIFPLKRECRVVLVEG 1685  
1681 KARGCAYPAPYLDYGETDPLGRGNPLHLSRERYEKLHVMQHCIIIEIARSQETNQ 1740  
1686 KARGCAYPAPYLDYGETDPLGRGNPLHLSRERYEKLHVMQHCIIIEIARSQETNQ 1745  
1741 LFGFNWQLL 1749  
1726 LFGFNWQLL 1734

RESULT 7  
US-10-758-636A-19  
; Sequence 19, Application US/10758636A  
; Publication No. US20050089876A1

GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966C  
; CURRENT APPLICATION NUMBER: US/10/758,636A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 19  
; LENGTH: 1734  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-758-636A-19  
  
Query Match 98.5%; Score 9087; DB 17; Length 1734;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1728; Conservative 0; Mismatches 1; Indels 20; Gaps 1;  
  
QY 1 MADBEAGTERMEISABELPQTPQRLASWDDQVDFTAFHLHLAQLVPEIYFAEMDPDLE 60  
DB 6 MADBEAGTERMEISABELPQTPQRLASWDDQVDFTAFHLHLAQLVPEIYFAEMDPDLE 65  
QY 61 KQESVQMSIPTLEWYLFGEEDPDI CLEKLKHSQAFOLCGRVPKSGETTYSCRDCAIDPT 120  
DB 66 KQESVQMSIPTLEWYLFGEEDPDI CLEKLKHSQAFOLCGRVPKSGETTYSCRDCAIDPT 125  
QY 121 CVLQWDCFDQSVHKNHRYKMHSTGGGFCDCGDTTEAMWKTGPFVYNHBPGRAGTIKENSRC 180  
DB 126 CVLQWDCFDQSVHKNHRYKMHSTGGGFCDCGDTTEAMWKTGPFVYNHBPGRAGTIKENSRC 185  
QY 181 PLNBEVIVQARKIPPSVIKYVVENTIWEBEKELPPELQIREKNRYCYVLFNDEHSHSYDH 240  
DB 186 PLNBEVIVQARKIPPSVIKYVVENTIWEBEKELPPELQIREKNRYCYVLFNDEHSHSYDH 245  
QY 241 VIYSIQALDCELAEALHTTAIDKEGRRAVKAGAYACQAEKEDIKSHSENVSOHPLHV 300  
DB 246 VIYSIQALDCELAEALHTTAIDKEGRRAVKAGAYACQAEKEDIKSHSENVSOHPLHV 305  
QY 301 EVLHSEIMAHQKFAIRLGLSNNKIMSYSSDFRQIFCQACLREBPDSNPCLISRLMLWDA 360  
DB 306 EVLHSEIMAHQKFAIRLGLSNNKIMSYSSDFRQIFCQACLREBPDSNPCLISRLMLWDA 365  
QY 361 KLYGARKILHELIFSSPFMEYKQLFAMEFVKYKQLOKEYISDDHDSISITALSVO 420  
DB 366 KLYGARKILHELIFSSPFMEYKQLFAMEFVKYKQLOKEYISDDHDSISITALSVO 425  
QY 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRIYAVICDLK 480  
DB 426 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRIYAVICDLK 485  
QY 481 YILISKPTIWTERRMQLEGFRSPLKILTQCMQGMEEIRROVGOHI EVDPDWEAIAIQM 540  
DB 486 YILISKPTIWTERRMQLEGFRSPLKILTQCMQGMEEIRROVGOHI EVDPDWEAIAIQM 545  
QY 541 QLKNTLMLFQEWACDEBLLLVAYKECHAVMRCTSTFISSTKVQSCGHSLETKSYRV 600  
DB 546 QLKNTLMLFQEWACDEBLLLVAYKECHAVMRCTSTFISSTKVQSCGHSLETKSYRV 605  
QY 601 SEDLVSHLPLSRTLAGLHVRLSRIGAVSRHLHFVSPEDFQVEVLVEYPLRCLVLVAQV 660  
DB 606 SEDLVSHLPLSRTLAGLHVRLSRIGAVSRHLHFVSPEDFQVEVLVEYPLRCLVLVAQV 665  
QY 661 AEMWERNGLSLISQVFFYQDVKCEEMVDKIIIMLQIGASLMDPNKFLLLVLQRYELAE 720  
DB 666 AEMWERNGLSLISQVFFYQDVKCEEMVDKIIIMLQIGASLMDPNKFLLLVLQRYELAE 725  
QY 721 FNKTIKTDQDLIKOYNTLIEEMQLVLIYVGERVYVGVGVNTKEEVTMRBIIHLLCIEP 780  
DB 726 FNKTIKTDQDLIKOYNTLIEEMQLVLIYVGERVYVGVGVNTKEEVTMRBIIHLLCIEP 785





Db 721 FNKTTSTKQDQLIKQNTLIEBMLQVLIYIVGERVYVGVNVTREBVMREITHLLCIEP 780  
QY 781 MPHSAIAKVLPEVNNETGLENNVINKVATFKKPGVSGHGVYELKDSESLKDFNMYFYHSK 840  
Db 781 MPHSAIARLPEVNNETGLENNVINKVATFKKPGVSGHGVYELKDSESLKDFNMYFYHSK 840  
QY 841 TQHSKAEHMKQRRQENKDEALPPPPPPFPAPSKVINLANCDIMMYILRTVPERAID 900  
Db 841 TQHSKAEHMKQRRQENKDEALPPPPPPFPAPSKVINLANCDIMMYILRTVPERAID 900  
QY 901 TDSNLWTEGMLQWAPHILALGALLEEKQLOKAPREBEVTFDYHKAASRLGSSAM---NTQM 957  
Db 901 TESNLWTEGMLQWAPHILALGALLEEKQLOKAPREBEVTFDYHKAASRLGSSAMNAQNIQM 960  
QY 958 LLEKLGIPOLGOKDMITWILQMDFTVKRLREKSLIVATTSGSEIKNDEITHDKKA 1017  
Db 961 LLEKLGIPOLGOKDMITWILQMDFTVKRLREKSLIVATTSGSEIKNDEITHDKKA 1020  
QY 1018 ERKRKAARLHRQIKMAQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEBESTPAVS DY 1077  
Db 1021 ERKRKAARLHRQIKMAQMSALQKNFIETHKLMYDNTSEVPKGEDSIMEBESTPAVS EA 1080  
QY 1078 SRIALGPKRGPSVTEKEVITCILCOBEQEVKTIENNAWVLSACVQKSTALTQHRGKPIEL 1137  
Db 1081 SRIALGPKRGPAVTEKEVITCILCOBEQEVKTIENNAWVLSACVQKSTALTQHRGKPV D HL 1140  
QY 1138 GHALDPLFMDPLAYGTYTSGCHVHVAWCWKYFEAVQLSSQORIHVDLPDLESEYLC 1197  
Db 1141 GETDPLFMDPLAHGTYTSGCHVHVAWCWKYFEAVQLSSQORIHVDLPDLESEYLC 1200  
QY 1198 PLCKSICNTVPIPILOPKINSADALAQLLTLARWLOTVLARTSGVNIHAKGENP- 1256  
Db 1201 PLCKSLCNTVPIPILOPKINSADALAQLLTLARWLOTVLARTSGVNIHAKGEAPA 1260  
QY 1257 IPIFFNQMGDSTLEPHSILSGVESSIKYSIKEMVILFATTIYRIGLKVPPDERDPR 1316  
Db 1261 VPVLFNQMGDSTLEPHSILSGVQSSVKYSIKEMVILFATTIYRIGLKVPPDELDPR 1320  
QY 1317 VPMLTWSTCAFTIQAENLLGDEGKPLFGALQNRQNGIKALMOPAVAOIRTCPOVLIOK 1376  
Db 1321 VPMTWTSTCAFTIQAENLLGDEGKPLFGALQNRQNGIKALMOPAVAOIRTCPOVL I HK 1380  
QY 1377 HLVRLLSVVLPNTKSEDTPLCLSIDLPHVLGAVLAPPSLYWDDPVDLPSPSVSSYNHL 1436  
Db 1381 HLARLUSVILPNLQSENTEPGLLSVDLPHVLGAVLAPPSLYWDDTVDLPSPSSSYNHL 1440  
QY 1437 YLFHLITMAHMLQILLTVDT----GLPLAQVOEDSEEAHSASSFFAIEISQYTSIGCDI 1492  
Db 1441 YLFHLITMAHMLQILLTTDLSLPGPLAEGBEDSEARCAAFFVEVSQHTDGLTGCGA 1500  
QY 1493 PGWYLWVSLKNGITPYLRCAALPHYVLLGVTPPEELHTNSAGEYSALCSYLSLPTNLF 1552  
Db 1501 PGWYLSLRLNGITPYLRCAALPHYVLLGVTPPEELFANSAGEFSALCSYLSLPTNLF 1560  
QY 1553 LFQEYWDTRPLLQRCADPALNCLKQNTVRYPRKNSLIELEPDDYSCILLNOASHPR 1612  
Db 1561 LFQEYWDTRPLLQRCWGGPALLKSLUKSAVVRYPRKNSLIELEPDDYSCILLNOASHPR 1620  
QY 1613 CPRSADDERKHPVLCIFCGAILCSQNICCOEIVNGEVEGACIFPHALHCGAGVCIFPKIRE 1672  
Db 1621 CPRSADDERKHPVLCIFCGAILCSQNICCOEIVNGEVEGACVPHALHCGAGVCIFPKIRE 1680  
QY 1673 CRVVLVEGARGCAYAPYLDYGETDPLKRGNPPLHLSRERYKRLHLVWQHCIIIEBIA 1732  
Db 1681 CRVVLVEGARGCAYAPYLDYGETDPLKRGNPPLHLSRERYKRLHLVWQHCIIIEBIA 1740  
QY 1733 RSQETNQLPFGFNWQLL 1749  
Db 1741 RSQETNQLPFGFNWQLL 1757

RESULT 9

US-10-758-636A-15

; Sequence 15, Application US/10758636A  
; Publication No. US20050089876A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966C  
; CURRENT APPLICATION NUMBER: US/10/758,636A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 15  
; LENGTH: 1757  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-758-636A-15

Query Match 93.4%; Score 861.7; DB 17; Length 1757;  
Best Local Similarity 92.1%; Pred. No. 0;  
Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;  
QY 1 MADBEAGGTGTERMEIGAEPLPQTPQRLASWMDQQVDFTTAFLLHHLAQLVPEIYFAEMDPDLE 60  
Db 1 MADEMDGAEKMDVSPPEPLAPQRPASWMDQQVDFTTAFLLHHLAQLVPEIYFAEMDPDLE 60  
QY 61 KOESVQMSIFTPPLWYLFGEDDDICLKLKHSGAFQICGRVFKSGGETTYSRCDCAIDPT 120  
Db 61 KOESVQMSILTPLEWYLFGEDDDICLKLKHSGAFQICGRVFKSGGETTYSRCDCAIDPT 120  
QY 121 CVLCHDCQDSVHKHRYKMTSTGGFCDCGDTAEWTKGPPCVNHEPRAGCTIKENSRC 180  
Db 121 CVLCHDCQSSVHKHRYKMTSTGGFCDCGDTAEWTKGPPCVNHEPRAGTTKESLHC 180  
QY 181 PLNBEVIVQARKIPPSVKYVEMTIWBEKEKLPPELQIREKNERYCYVLFNDEHHSYDH 240  
Db 181 PLNBEVIAQARRIPPSVKYVEMTIWBEKEKLPPELQIREKNERYCYVLFNDEHHSYDH 240  
QY 241 VIYSIQRALDCELABAQLHTTATIDKEGRRVAVAGAYACQEAKEKDIKSHSENVSHPLHV 300  
Db 241 VIYSIQRALDCELABAQLHTTATIDKEGRRVAVAGAYATCQEAKEKDIKSHSENVSHPLHV 300  
QY 301 EYLHSEIWAHQFALRIGSWNKNKINSYSDPQIFCOACLRPEPDSNPCLISRLMLWDA 360  
Db 301 EYLHSEVMAHQFALRIGSWNKNKINSYSDPQIFCOACLRPEPDSNPCLISRLMLWDA 360  
QY 361 KLYKGARKILHELIFSSFFMEMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALS VQ 420  
Db 361 KLYKGARKILHELIFSSFFMEMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALS VQ 420  
QY 421 MFTVPTLARHLTEBQNVISVITETLLEVLPEYLDNNKNFNFGYSQDKLGRVAVICDLK 480  
Db 421 MLTVPTLARHLTEBQNVISVITETLLEVLPEYLDNNKNFNFGYSQDKLGRVAVICDLK 480  
QY 481 YLILSKPTIWTBRLAQFLGERSFLKILTCMQGMEIIRROVGOHLEVPDDEAAIAIOM 540  
Db 481 YLILSKPVIWTBRLAQFLGERSFLKILTCMQGMEIIRROVGOHLEVPDDEAAIAIOM 540  
QY 541 QLNKILLAFQEWACDEBELLVAYKECHKAVNRCSTSFISSTKTVVQSCGHSLETYSYV 600  
Db 541 QLNKILLAFQEWACDEBELLVAYKECHKAVNRCSTINFMSSTKTVVQSCGHSLETYSYV 600  
QY 601 SEDLVSIHLPLSRTLAGLHVLSRLGAVSRHLHFVSFDFQVEVLVEYPLCLVLVAQV 660  
Db 601 SEDLVSIHLPLSRTLAGLHVLSRLGAVSRHLHFVSFDFQVEVLVEYPLCLVLVAQV 660  
QY 661 AEWWRNGLSLSQVFFYQDVVKCREMDKDIIMLQIGASLMDPNKFLLLVLQRYELAE 720  
Db 661 AEWWRNGLSLSQVFFYQDVVKCREMDKDIIMLQIGASIMDPNPKFLLLVLQRYELTDA 720  
QY 721 FNKTTSTKQDQLIKQNTLIEBMLQVLIYIVGERVYVGVNVTREBVMREITHLLCIEP 780



Db 721 FNKTIISKQDLIKQNTLIEEMQLVLIYVGRYVPGVGNVTRREVMRRETHLJCIEP 780  
QY 781 MPHSALAKNLPENENNETGLENVINKVATPKPGVSGHGVYELKDBSLKDFNFPYFHYSK 840  
Db 781 MPHSALAKNLPENENNETGLENVINKVATPKPGVSGHGVYELKDBSLKDFNFPYFHYSK 840  
QY 841 TQSKAHEHQKRRKQENKDEALPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 900  
Db 841 TQSKAHEHQKRRKQENKDEALPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 900  
QY 901 TDSNLTGEMQWAFHILALGLLEKQKQKAPPEEVTDFVHKASRLGSSAM---NIOM 957  
Db 901 TESNLTGEMQWAFHILALGLLEKQKQKAPPEEVTDFVHKASRLGSSAMNAQIQM 960  
QY 958 LLEKLAGIPQLEGQKDMITWILQMDTVKRLREKSLIVATTSGSBSIKNDIETHDKKA 1017  
Db 961 LLEKLAGIPQLEGQKDMITWILQMDTVKRLREKSLIVATTSGSBSIKNDIETHDKKA 1020  
QY 1018 ERKKAARLHRQKIMQASALQKPIETHKLMYDNTSEMPKQKDSIMEESTPAVS DY 1077  
Db 1021 ERKKAARLHRQKIMQASALQKPIETHKLMYDNTSEMPKQKDSIMEESTPAVS DY 1080  
QY 1078 SRIALGPKGSPVTEKEVLTICLCOEQEVKINNANWLSACVOKSTALTQHRGKPIELS 1137  
Db 1081 SRIALGPKGSPVTEKEVLTICLCOEQEVKINNANWLSACVOKSTALTQHRGKPIELS 1140  
QY 1138 GBALDPLFMDPOLAYTYTGSCHVNHAVCWQYFEAVOLSSOORLHVDFLSEGEYLC 1197  
Db 1141 GETLDPLFMDPOLAYTYTGSCHVNHAVCWQYFEAVOLSSOORLHVDFLSEGEYLC 1200  
QY 1198 PLCKSLCNTVPIPILOPQKINSADALALQTLTARWOTVLARTSGYNIRHAKGPN- 1256  
Db 1201 PLCKSLCNTVPIPILOPQKINSADALALQTLTARWOTVLARTSGYNIRHAKGPN- 1260  
QY 1257 IPFPNOGQDSTLEPHSILSGVSSIKYNSIKEMVILPATTIYRIGLKVPPDERDR 1316  
Db 1261 VPVLPNOGQDSTLEPHSILSGVSSIKYNSIKEMVILPATTIYRIGLKVPPDERDR 1320  
QY 1317 VPMLTWTSTCAFTIQAENLIDGEGKPLFGALQNRQNGIKALQWAPAVARITCPOVLIOK 1376  
Db 1321 VPMTWTSTCAFTIQAENLIDGEGKPLFGALQNRQNGIKALQWAPAVARITCPOVLIOK 1380  
QY 1377 HLVRLLSVVLPNIKSDPTCLLSIDLPHVLGAVLAFPSLYWDDPVDLPSPVSSSYNHL 1436  
Db 1381 HLARLLSVLPNIKSDPTCLLSIDLPHVLGAVLAFPSLYWDDPVDLPSPVSSSYNHL 1440  
QY 1437 YLFHLITMAHMLQILITVDT---GLPLAQVEDSEAHASFPFABISQYTSIGCDI 1492  
Db 1441 YLFHLITMAHMLQILITVDT---GLPLAQVEDSEAHASFPFABISQYTSIGCDI 1500  
QY 1493 PGWYLVLSIKNGITPVLRCALPHYLIGVTPPELHTNSARGESYALCSYLSLPTNLP 1552  
Db 1501 PGWYLVLSIKNGITPVLRCALPHYLIGVTPPELHTNSARGESYALCSYLSLPTNLP 1560  
QY 1553 LFQEWYDVTVRLLQRCADPALLNCLIKQNTVVRYPKNSILIELPDDYSCILNOASHPR 1612  
Db 1561 LFQEWYDVTVRLLQRCADPALLNCLIKQNTVVRYPKNSILIELPDDYSCILNOASHPR 1620  
QY 1613 CPRSADDERKHPVLCFCGAILCSQNICQEIIVNGEVEGACIFHALHCGAGVCIPLKIRE 1672  
Db 1621 CPRSADDERKHPVLCFCGAILCSQNICQEIIVNGEVEGACIFHALHCGAGVCIPLKIRE 1680  
QY 1673 CRVNLVEGKARGCAPAPVLYDEYGETDPLGKGNPLHLSREYRKHLVWQOHCIIIEBTA 1732  
Db 1681 CRVNLVEGKARGCAPAPVLYDEYGETDPLGKGNPLHLSREYRKHLVWQOHCIIIEBTA 1740  
QY 1733 RSQETNOMLFGFNWQLL 1749  
Db 1741 RSQETNOMLFGFNWQLL 1757

US-10-758-672A-4  
; Sequence 4, Application US/10758672A  
; Publication No. US20040185037A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.  
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966B  
; CURRENT APPLICATION NUMBER: US/10/758,672A  
; PRIOR FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US 09/724,126  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 4  
; LENGTH: 1755  
; TYPE: PRT  
; ORGANISM: Homo sapeins  
; US-10-758-672A-4

Query Match 46.9%; Score 4327; DB 16; Length 1755;  
Best Local Similarity 47.2%; Pred. No. 0;  
Matches 841; Conservative 339; Mismatches 543; Indels 58; Gaps 27;

QY 1 MADE---EAGGTER--MEISAEIPQTPQRLASWWDQVDFYTAFLHLAQLVPEIYFAEM 55  
Db 1 MASELEPEVQADRSILECSAE-----EIAKGMLQATDLTREVYQHLAHYVPKIYCRGP 54  
QY 56 DPDLSEKQESVQMSIPTLEWYLFQEDDPDICLEKUKHSG-AFQLCGRVPKSGSETTYSCRD 114  
Db 55 NFPPOKEDMLAQHVLGLPMEWYLCGEDPAFGPKLEQANKPSHLGRVFKVGEPTYSCRD 114  
QY 115 CAIDPTCVLCMDQFQDSVHKHRYKMHSTGTGGFCDCGDEAWKTGPFVNVHEPGRAG-T 173  
Db 115 CAVDPTCVLCMECFGLSHRDHRYRMTTSGGGFCDCGDEAWKEGFCCKHELNTSEIE 174  
QY 174 IKENSRCPLNEVIVQARKIPFVIVVEMTWEEKELPPELOIREKNERYCVLFFND 233  
Db 175 BEEDPLVHLSVEDVIARTYNIPTATRYAVAILTWEESELPADLENVEKSDTYCMLFND 234  
QY 234 EHSYDHYIYSIQRALDCELAQAHLHTTAIDKEGRRVAKAGAYAAQCEAKEDIKHSENV 293  
Db 235 EVHTYEQVITYTLQKAVNCTQKEAIGFATTVDRGRSRYGDFQYCEQAKSVIVRNTSRQ 294  
QY 294 SQHPLHVEVLHSEIMAHQKFAIRLGLSNWKNKINSYSDPRQIFQACLRSEPSSENPLIS 353  
Db 295 TK-PLKVQVMHSSIHAHQNFGLKLSWLSIIGYSDGLRRLCQVGLQSGPGGNSSLVD 353  
QY 354 RLMLMDAKLYKGARKILHELIPSPFMEBYKKLFAEMBYKVKYKOLQKEYISDDHDSRIS 413  
Db 354 RLMLSDSKLWKGARSVYHQLFMSLLMDLYKYLKLVAVRPAKQYQQLQDPMEDDHERAVS 413  
QY 414 ITALSVQMTVPTLARHLIHEONVISVITETLLEVLPEYLDNRNKNFQGIS---QDKLG 470  
Db 414 VTALSVQFTAPTARMLITEENLSIIKTFMDHL-RHRDAQGRFQFERYTALQAFKER 472  
QY 471 RYAVICDLYLISKPTIWERLBMQFLEGRSFLKILTQMGHEERTRQVGOHIEVDP 530  
Db 473 RVQSLILDKYVLSKPTENSDELQKFLQEGDPAFLLELLKCMQGMDPITRQVGOHIEEP 532  
QY 531 DWEAIAITOMOLKNITLLMPQEWACDBELLVAYKSKCHAVMRCSTSFSSSKTVQS-C 589  
Db 533 EWEAFTTOMKLTHTVISHMQDWACDEKVLIEAYKKCLAVLMQCHGGYDGDGQPTLSLC 592  
QY 590 GHSLETKSYRVEDLSVSIHLPLSRTLGLVRLSLGAVSRHLHEFVSPEDFOVEVLVEYP 649  
Db 593 GHSVETIRYVSQEKVSIHLPVSRLLAGLHVLSKSEVAYKPELPLSELSPPLIEHP 652  
QY 650 LRCLVLVAQVAVMWRNGLSLISQVFTYQDVQCKREMYDDIIMLOITGASLMDNPKPL 709  
Db 653 LRCLVLCAQVHAGMWRNGLSLVNIQIYTHVYKCRREMFDKOVVLMQTVGSMMDPNHFLM 712







653 LRCLVLCQVHAGWRNGESLVNQYVYHNKCRREMFQKDIYMLQTVGSMMDPNHFLM 712  
710 LVLRQVELAFN-----KTIATK--DQDLIKQYNTLIEBMLQVLIYVGERVYVPGVNV 762  
713 IMLSRPELYQLFSTPDYKGRFSSEVTHKDVVQNNLIEBMLYLIIMLVGERPNPGVQV 772  
763 TKEEVMTRETIHLLCTEPHPSAIAKNNPENNETGLENVINKVATPKKPGVSGHVYE 822  
773 AATDEIKREITHQLSIKPMAHSELVSLPDEKNTGMSVIEBSVAHFKKPGITGRMTE 832  
823 LKDESILKDFMYFYHYSKTOHSAEHMQKRRQENKDEALPPPPPEFCFAPSKVINLL 882  
833 LKPECAKBNLYYHFSRAEQSKAEARQKLRKREKEDTALPPALPPFCPLFASLVNLL 892  
883 NCDIMMYLRTVPERAIDTDSNLWTEGMLQMAFHIALGLLEKQOLQKAPEREV-TFDF 941  
893 QCDVLYIMGTILQWAVEHNGSAWSEMLQVRLHGLMALQEBKHLENAVEGHVQTFTF 952  
942 YHKASRLGSSAMN---IQMLEKLGKIPQLEGOKDMITWILQVDTVKRLREKSLIVAT 998  
953 TQKISKPGDAPHNSPILAMLETQWAPSLAEAKDMIRMLKMFNAIKIRE--CSSSSP 1010  
999 TSGSESIKNDIETHDKAEKREKAEARLHROKIMAOHSAKQKNEIETHKLMYDNTSEM 1058  
1011 VAEAGTIEESRDKDAERKGAIAIRLRKIMAOHSEMRHFDENKELFOOTLEL 1070  
1059 PGKEDSIMBESSTPAVDYSRIALGPKRGSVTEKEVLTCLCOEQEYKINNANVLISA 1118  
1071 DTSASATL--DSSPPVSDAALTALGPAQTVQVPEPRQFVTCILCOEQEYVTSRANVLAA 1128  
1119 CVOKSTALQHRGKPIELSGEALDPLFMDPLIAYTYTSCGHHMAVACQKYEAVQLS 1178  
1129 FVORSTVLKORTKIT-ADPEKYDPLFHPDLSCGTHTSCGHHMAHCAWQYFDSVQAK 1187  
1179 SQ-----RIHVDLPLEGEVLCPLCKSICNTVPIIPLQPKINSENADALAQLLTIA 1233  
1188 EQRORQLKHTS-YDVNGEFLCPLCECLSNVTPIPL-LPPRSILSRRLN-FSDQPDIA 1244  
1234 RWTQTLARISGNYI---RHAKGEN-----PIPIFNQMGSDSTLEFHSILSF 1278  
1245 QMTRAVTQIKVQVQLRKHNAADTSSSEDTAMNIIPIPEGRP-----DFY----- 1292  
1279 GVESSIKYSNSIKEMVILPATTYIRIGLVKPPDERDPRVPMLTWSTCAFTIOAENLGD 1338  
1293 ---PRNPYSDSIKEMLTTFGTAAKVKLVHNEGDPVPLICWGTGCAVTIQSIERILSD 1349  
1339 ECKPLFEGALQNRHGLKALMOPAVARITCQVLIQKHLVRLSVLVNPKISBDTPCLL 1398  
1350 EKPVPFGLPCRLDDCLRSLTRFAAAHVTVALPVVQGHFCKLFASLVPBDSYEDLPCIL 1409  
1399 SIDLPHVLGAVLAPPSLYWDDPVDLPQSSVSSYNHLYLPHLITMAHMLQILL---TVD 1455  
1410 DIDMFHLLVGLVLAFLAQCDQ-----FSGSSLATGDLHIFHLVTHAHTVQILLTSCTEB 1464  
1456 TGLPLAQVEDSEEAHSAASSPAETISQYTSIGICDIPGWLWVLSKNGITPYLCAALF 1515  
1465 NGMD--QENPTGEBELIILSLKTLHQYTGSAKCAPSCHLWRSVRAIMPFKCSALF 1522  
1516 PHYLGVTPPEBLHNSABGEYSALCSYLSLPTNIFLLPQEWYDTRVPLLRWCADPALL 1575  
1523 PHYLNGVPAPDQLV-SGTSHPEHLNLYSLPTNIFLLPQENSIDMNSLIESWCQNSEVK 1581  
1576 NCLKQKTVVYPRKENSILIELPDDYSCLLNOASHPCRSADDERKHPVLCFGALIC 1635  
1582 RYLNGERGAISYPRGANKLIDLPEDYSSLIQNSAFSCPKSGDKSRAPTLCLVCGSLIC 1641  
1636 SONICCOETVNGEEVACIFHALHCGAGVCIPLKIRECRVVLVEGKARGCAYAPYLDY 1695  
1642 SOSYCCQAELEGEDVGACTAHTYSCTSGNGIFLRVRECVQLFAGTKGCTFSPYLDY 1701  
1696 GETDPGLKRGNPLHLSRERYRKLHLVWQOHCIIIEBARSQETNQMFLGPNWOLL 1749

Db 1702 GETDQGLRGNPLHLCOERFKRKLQKLMQOQHSITBIGHAQBANQTLVGIDMQLH 1755  
RESULT 14  
US-10-755-889-635  
; Sequence 635, Application US/10755889  
; Publication No. US20040171823A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
; TITLE OF INVENTION: PATHWAY  
; FILE REFERENCE: D0284 NP  
; CURRENT APPLICATION NUMBER: US/10/755,889  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
; PRIOR FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: U.S. 60/469,757  
; PRIOR FILING DATE: 2003-05-12  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 635  
; LENGTH: 1275  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-755-889-635  
Query Match 34.3%; Score 3160; DB 16; Length 1275;  
Best Local Similarity 47.3%; Pred. No. 1.7e-250;  
Matches 612; Conservative 248; Mismatches 393; Indels 40; Gaps 19;  
QY 479 LKYLISKPTIWTBRLRQFLEGFRSFLKILCMQGBEIRQVQGHIEVDPDWRAAIAI 538  
Db 1 LKYLISKPTIWTBRLRQFLEGFRSFLKILCMQGBEIRQVQGHIEVDPDWRAAIAI 538  
QY 539 QMLAKNILLPQEWCADEELLVAYKCHKAVMCSTSFISSTKVVOG-CGHSLETYS 597  
Db 61 QMLKTHVISMDDWCASDEKVLIEAYKKCLAVMLQCHGYTDEGRPTILSICHSVETTR 120  
QY 598 YRVSBDLVSIHLPLSRTLAGLHVRLSLGAVSRLEHFEVDFQVEVLVEYPLCLVLVA 657  
Db 121 YCVSEKYSIHLPVSRLLAGLHVLSKSEVAYKFPELLPLSELSPMLIEHPLRCLVLA 180  
QY 658 QVVAEMRRNGLSISQVYFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLRQYEL 717  
Db 181 QVHAGWRNGESLVNQYVYHNKCRREMFQKDIYMLQTVGSMMDPNHFLMILSRPEL 240  
QY 718 AEAFF-----KTIATK--DQDLIKQYNTLIEBMLQVLIYVGERVYVPGVNTTKEVTHR 770  
Db 241 YQIFSTPDYKGRFSSEIETHKDVVQNNLIEBMLYLIIMLVGERFSPGVQVNAIDIKR 300  
QY 771 ETLHLLCTEPHPSAIAKNNPENNETGLENVINKVATPKKPGVSGHVYELKDESLKD 830  
Db 301 ETLHLLCTEPHPSAIAKNNPENNETGLENVINKVATPKKPGVSGHVYELKDESLKD 360  
QY 831 FNMVYFYSKTOHSAEHMQKRRQENKDEALPPPPPEFCFAPSKVINLNCIMMYI 890  
Db 361 FNMVYFYSKTOHSAEHMQKRRQENKDEALPPPPPEFCFAPSKVINLNCIMMYI 420  
QY 891 LRTVPERAIDTDSNLWTEGMLQMAFHIALGLLEKQOLQKAPEREV-TFDFYHKAIRIG 949  
Db 421 MGTILQWAVEHNGSAWSEMLQVRLHGLMALQEBKHLENAVEGHVQTFTFQKISKPG 480  
QY 950 SSAMN---IQMLEKLGKIPQLEGOKDMITWILQVDTVKRLREKSLIVAT 1005  
Db 481 EAPKNSPILAMLETQWAPSLAEAKDMIRMLKMFNAIKIRE--SPALETGTTM- 539  
QY 1006 KNDIETHDKAEKREKAEARLHROKIMAOHSAKQKNEIETHKLMYDNTSEMPCEDSI 1065  
Db 540 --BESSRDKDAERKGAIAIRLRKIMAOHSEMRHFDENKELFOOTLELDASTAV 597  
QY 1066 MEEESTPAVDYSRIALGPKRGSVTEKEVLTCLCOEQEYKINNANVLISACVQKSTA 1125  
Db 598 LDH--SPVASDWTLTALGPTQTQVPEQRQFVTCILCOEQEYKINNANVLISACVQKSTA 655

```

QY 1126 LTOHRKDIELSSEALDPLFMDPDLAYTYTSCGCHVHAYCQKYFEAVQLSSQ----- 1181
DB 656 LSKRSKFIQ-DPEKYDFLFHMDPLSCGTHTSSCGHIMHAWCQRYFDFVQAKGQROR 714
QY 1182 -RIHVDLDBLSGEYCLPCLSLCNTVPIPILOPKINSNADALAQLLTLARWITVL 1240
DB 715 LRLHTS-YDVENGEPFLPCECLNTVPIPL-LPRNLFNNRLN-PSDPNLTQWIRIS 771
QY 1241 ARISGYNIRHAKENPIPIFNQGMGSTLBFHSILSPGVSSIKYISIKEMVILFATT 1300
DB 772 QQIKALQFLRKEESTPNNASTKNSNVDELQLPFGFRDPFRPKIPYSSIKEMLTFTFTA 831
QY 1301 IYRIGLVKPPDRDRPVMLTWSCTAFTIOAENLLGEGPLFGLALQNRQHGLKALMQ 1360
DB 832 TYKVLKRVHNEEDPRVPMCGSCAYTIOSTIERILSDKPLFGPLPCRLDDCLRLSUTR 891
QY 1361 PAVAQRTICPOVLIOKHLVRLSVVLPNPKISDTPCLLSIDLPHVLGAVLAFPSLYWDD 1420
DB 892 FAHAHTVNASVSVQCHFKLPASLVFNDSSHEELPCILIDMFHLLVGLVLAFFALQCOD 951
QY 1421 PVDLPSSVSSNHYLFLHITMAHMLQILL---TVDTGLPLAQVQSDSEAHASGFF 1477
DB 952 -----FSGISLGTGDLHIFLHVTMAHIIQILLTSCTEENGMD--QENPPCEESAVALAY 1004
QY 1478 ABISQYTSIGCDIP-GWYLVSLKNGITPILRCAALFFHYLLGVTPPEELHTNSAGE 1536
DB 1005 KTLHQYT-GSALKETPSGHLWRSVRAGIMPLFKCSALFFHYLNGVPSPDDIQV-PGTSH 1062
QY 1537 YSALCSYLSPLTNLFLPQXYWDTVRPLLORCADPALLNCLKOKNTVVRYPRKNSLIE 1596
DB 1063 PHLCSYLSPLNKLCLFQENSEIWNLSIESCRNVRVLEGERRDAIRYPRESKNLIN 1122
QY 1597 LPDDYSCILNQASHFRPSADDERKHPVLCLFCGAILCSQNICQBIVNGEVCACIFH 1656
DB 1123 LPEDYSSILNQASHNFCPSKSGDKSRAPTLCLVCSGLLCSQYCCQTELEGEDVGACTAH 1182
QY 1657 ALHCGAGVICFLKIRECRVVLVEGKARGCAYPAPVDEYGETDGLKRGNPLHLSRYR 1716
DB 1183 TYSCSGVGIFLVRRECQVLFLAGTKGCFYSPYLDYDGETDOGLRRGNPLHLCKERFK 1242
QY 1717 KLHLVWQHCITEELARSOETNQLFGFNWQLL 1749
DB 1243 KIQLWQHQSVEEIGHAQEANTLVIGIDWQHL 1275

```

RESULT 15

```

US-10-094-749-2555
; Sequence 2555, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094, 749
; CURRENT FILING DATE: 2002-03-12

```

```

; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2555
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2555

Query Match      33.2%; Score 3066; DB 15; Length 594;
Best Local Similarity 100.0%; Pred.No. 2.8e-243;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 MFTVPTLARHLIERQNVISVITETLLEVLPRYLDRNNKFNQGSODKLGRIYAVICDLK 480
DB 1 MFTVPTLARHLIERQNVISVITETLLEVLPRYLDRNNKFNQGSODKLGRIYAVICDLK 60
QY 481 YILISKPTIWTERRMOLFEGFRSLKILTCMQGMEIRROVGOHIEVDPDWEAIAIQM 540
DB 61 YILISKPTIWTERRMOLFEGFRSLKILTCMQGMEIRROVGOHIEVDPDWEAIAIQM 120
QY 541 QLKNIILLMFQWCACDELLIVAYKECHKAYMRCSTSFISSTKVQSCGHSLETKSYRV 600
DB 121 QLKNIILLMFQWCACDELLIVAYKECHKAYMRCSTSFISSTKVQSCGHSLETKSYRV 180
QY 601 SEDLVSIHLPSRLTAGLHVRLSRLGAVSRHLHFVSPEDQVTVLVYVPLCLVLVAQVV 660
DB 181 SEDLVSIHLPSRLTAGLHVRLSRLGAVSRHLHFVSPEDQVTVLVYVPLCLVLVAQVV 240
QY 661 AEMWRRNGLSLISQVFFYYQDVKCEEMVYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 720
DB 241 AEMWRRNGLSLISQVFFYYQDVKCEEMVYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 300
QY 721 FNKTIISTKQDILLIKQYNTLIBEMLQVLIYIVGERYVPGVGNVTKEEVTMRIIHLLECIEP 780
DB 301 FNKTIISTKQDILLIKQYNTLIBEMLQVLIYIVGERYVPGVGNVTKEEVTMRIIHLLECIEP 360
QY 781 MPHSAIAKNLPENENNETGLENVINKVATPKKPGVSGHGVYELKDESUKDFNMYPHYYSK 840
DB 361 MPHSAIAKNLPENENNETGLENVINKVATPKKPGVSGHGVYELKDESUKDFNMYPHYYSK 420
QY 841 TQHSKAEHMOKKRRQENKDEALPPPPPPPCPAPSKVINLINCDDIMVILRTVPERAID 900
DB 421 TQHSKAEHMOKKRRQENKDEALPPPPPPPCPAPSKVINLINCDDIMVILRTVPERAID 480
QY 901 TDSNLWTEGMLQMAFHIALGILLBEKQOLQKAPEEVTFDFYHKASRLGSSAMNIQMLLE 960
DB 481 TDSNLWTEGMLQMAFHIALGILLBEKQOLQKAPEEVTFDFYHKASRLGSSAMNIQMLLE 540
QY 961 KLKGIPOLEGQKDMITWILQMPDVTVKRLREKSLIVATTSSESINKNDEITHDK 1014
DB 541 KLKGIPOLEGQKDMITWILQMPDVTVKRLREKSLIVATTSSESINKNDEITHDK 594

```

Search completed: July 9, 2005, 15:43:25  
Job time : 142 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 13:43:17 ; Search time 47 Seconds  
(without alignments)  
3580.492 Million cell updates/sec

Title: US-10-758-672A-2  
Perfect score: 9224  
Sequence: 1 MADEAGGTERMEISAEPLQ.....ETARSQETNQLFGFNWQLL 1749

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8617	93.4	1757	2 T14318	ubiquitin-protein
2	1945.5	21.1	1927	2 T25604	hypothetical prote
3	1038	11.3	1958	2 T39808	hypothetical prote
4	914	9.9	2052	2 T37711	probable n-end-rec
5	689.5	7.5	1950	2 S12332	ubiquitin-protein
6	647.5	7.0	1941	2 T30554	ubiquitin-protein
7	606.5	6.6	1225	2 T48251	ubiquitin-protein
8	549.5	6.0	1872	2 S64851	probable membrane
9	537.5	5.8	2168	2 D88131	protein F10G7.10 (
10	343	3.7	795	2 T48252	eceriferum3 (CERS)
11	218	2.4	271	2 T40238	hypothetical prote
12	205.5	2.2	3187	2 JCS837	364K Golgi complex
13	190.5	2.1	3259	1 A56539	giantin - human
14	186	2.0	1410	1 A57013	early endosome ant
15	182	2.0	2442	2 T08621	centrosome associa
16	180.5	2.0	1413	2 T28467	hypothetical prote
17	180	2.0	1830	1 S19188	myosin-v - chicken
18	179.5	1.9	2346	2 T13829	Tpr homolog - frui
19	176	1.9	3225	2 I52300	giantin - human
20	175.5	1.9	1427	2 S22695	restin - human
21	175.5	1.9	1526	2 T41522	myosin ii - fission
22	175.5	1.9	2863	1 S28261	centromere protein
23	174.5	1.9	1009	2 C89910	hypothetical prote
24	174.5	1.9	1392	2 A43336	microtubule-vesicl
25	172.5	1.9	1780	2 T17272	hypothetical prote
26	168	1.8	1837	2 T41023	probable nuclear p
27	166	1.8	1163	2 G97236	ATPase involved in
28	166	1.8	1676	2 E71410	probable centromer
29	164.5	1.8	1181	2 T01799	hypothetical prote

30	164	1.8	2954	2 T14156	kinesin-related pr
31	163.5	1.8	1339	2 H81307	restriction modifi
32	163	1.8	1853	1 A46761	myosin heavy chain
33	163	1.8	2649	2 A40937	bullous pemphigoid
34	162	1.8	1269	2 F84730	probable myosin he
35	157.5	1.7	1642	2 T08880	NMDA receptor-bind
36	156.5	1.7	1496	2 T05634	hypothetical prote
37	156.5	1.7	1790	2 S67593	transport protein
38	156.5	1.7	2710	2 A37052	reticulocyte-bind
39	156.5	1.7	2829	2 A42771	toxin A - Clostrid
40	155.5	1.7	1216	2 A28821	1-phosphatidylinos
41	155	1.7	1025	2 E86355	hypothetical prote
42	155	1.7	1173	1 A53430	1-phosphatidylinos
43	155	1.7	1828	2 B59254	myosin heavy chain
44	154.5	1.7	2269	2 T28677	rhopty protein -
45	154	1.7	3685	1 A27605	dystrophin, muscle

ALIGNMENTS

RESULT 1

T14318  
ubiquitin-protein ligase E3-alpha - mouse  
NAlternate names: N-recognin E3-alpha  
CSpecies: Mus musculus (house mouse)  
CDate: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
CAccession: T14318  
R:Kwon, Y.T.; Reiss, Y.; Fried, V.A.; Herahko, A.; Yoon, J.K.; Gonda, D.K.; Sanghan, P.;  
Proc. Natl. Acad. Sci. U.S.A. 95, 7898-7903, 1998  
ATitle: The mouse and human genes encoding the recognition component of the N-end rule  
AReference number: Z17977; MUID:98318583; PMID:9653112  
AAccession: T14318  
AStatus: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: mRNA  
A.Residues: 1-1757 <KNO>  
A.Cross-references: UNIPROT:070481; EMBL:AF061555; MID:g3170886; PID:g3170887; PIDN:AAC  
C.Genetics:  
A.Gene: Ubr1  
A.Map position: 2

Query Match	93.4%	Score	8617	DB	2	Length	1757
Best Local Similarity	92.1%	Pred. No.	0				
Matches	1619	Conservative	64	Mismatches	66	Indels	8
Gaps	3						
QY	1	MAD	EAGGTERMEISAEPLQTPQRLASWWDQVDFTAFLLHQAQLVPEIYFAEMDPDL	60			
DB	1	MAD	EMDGAERMDVSPPEPLAPQRPASWWDQVDFTAFLLHQAQLVPEIYFAEMDPDL	60			
QY	61	KQ	ESVQMSIFPTLEWYLFGEEDPDICLEKLKHSQAFQICGRVFKSGTYSRCDCAIDPT	120			
DB	61	KQ	ESVQMSILTPLEWYLFGEEDPDICLEKLKHSQAFQICGRVFKSGTYSRCDCAIDPT	120			
QY	121	CVL	CHDCQDSVHKHRYKMTSTGGCDGDTAWKTPPCVNHFPGRAGTIKENSRC	180			
DB	121	CVL	CHDCQSSVHKHRYKMTSTGGCDGDTAWKTPPCVNHFPGRAGTIKESLHC	180			
QY	181	PLN	EVIVQARKIPPSVIKYVEMTWIWEKKELPELOIREKNRYCYVLPNDEHSHVDH	240			
DB	181	PLN	EVIAQARKIPPSVIKYVEMTWIWEKKELPELOIREKNRYCYVLPNDEHSHVDH	240			
QY	241	VIY	LSQALDCELAQAQLHTTAIDKGRRAVKAGAYACQEKEDIKSHSENVSOHPLHV	300			
DB	241	VIY	LSQALDCELAQAQLHTTAIDKGRRAVKAGAYATCQEKEDIKSHSENVSOHPLHV	300			
QY	301	EVL	HSIWAHOKFALRLGSSWANKMSYSSDFRQIFCQACLRREPDSENPCLISRLMLWA	360			
DB	301	EVL	HSVWAHOKFALRLGSSWANKMSYSSDFRQIFCQACLRREPDSENPCLISRLMLWA	360			
QY	361	KLY	GARKILHELIFSSPFMEYKCLFAMSFVKYKQLOKEYISDDHDSRISITALS	420			
DB	361	KLY	GARKILHELIFSSPFMEYKCLFAMSFVKYKQLOKEYISDDHDSRISITALS	420			

QY 421 MFTVPTLARHLIEEQNVISVITETLLLEVLPEYLDNRNKENFQGYSDKLGRRVAVICDLK 480  
DB 421 MFTVPTLARHLIEEQNVISVITETLLLEVLPEYLDNRNKENFQGYSDKLGRRVAVICDLK 480  
QY 481 YILISKPTIWTBRLRMQFLGFRSFLKILTCMQGMBEIRROVGOHIEVDPDWEAAIAIQM 540  
DB 481 YILISKPTIWTBRLRMQFLGFRSFLKILTCMQGMBEIRROVGOHIEVDPDWEAAIAIQM 540  
QY 541 QLNKILLMFOEWACDEBELLVAYKECHKAVMRCSTSPISSTKTVVQSGCHSLKTSYRV 600  
DB 541 QLNKILLMFOEWACDEBELLVAYKECHKAVMRCSTSPISSTKTVVQSGCHSLKTSYRV 600  
QY 601 SEDLSIHLPLSRTLAGHVRSLGAVSLHFEVSPEDFOVEVLVEYPLCLVLVAQVV 660  
DB 601 SEDLSIHLPLSRTLAGHVRSLGAVSLHFEVSPEDFOVEVLVEYPLCLVLVAQVV 660  
QY 661 AEMWRNGLSLISQVFPYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 720  
DB 661 AEMWRNGLSLISQVFPYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 720  
QY 721 FNKTISTKQDOLIKQNTLIEEMQLVLIYVGERVYVGVNVTKEBVTWREIHLICIRP 780  
DB 721 FNKTISTKQDOLIKQNTLIEEMQLVLIYVGERVYVGVNVTKEBVTWREIHLICIRP 780  
QY 781 MHSATAKNLPENENNETGLEVINVKVATPKPGVSGHGVYELKDBSLKDFNMFVYHSK 840  
DB 781 MHSATAKNLPENENNETGLEVINVKVATPKPGVSGHGVYELKDBSLKDFNMFVYHSK 840  
QY 841 TOHSAKHAHQKRRQKQENKDEALPPPPPPPPFAPSKVINLNCIDIMVILRTVFERAID 900  
DB 841 TOHSAKHAHQKRRQKQENKDEALPPPPPPPPFAPSKVINLNCIDIMVILRTVFERAID 900  
QY 901 TOSNLWTEGMLQWAFHILALGLLEKQOLQKAPPEBVTDFYHKASRLGSSAM---NIQM 957  
DB 901 TOSNLWTEGMLQWAFHILALGLLEKQOLQKAPPEBVTDFYHKASRLGSSAMNAQIQM 960  
QY 958 LLEKLKGIPOLEQKQDMITWILQMPDFTVRLREKSLIVATTSGSBSIKNDRIETHDKEA 1017  
DB 958 LLEKLKGIPOLEQKQDMITWILQMPDFTVRLREKSLIVATTSGSBSIKNDRIETHDKEA 1020  
QY 1018 ERKKAABARHLRQKIMQMSALQKNFIETHKLYMDNTSEMPGKDSIMEEESTPAVSIDY 1077  
DB 1021 ERKKAABARHLRQKIMQMSALQKNFIETHKLYMDNTSEVTKGDSIMEEESTSAVSEA 1080  
QY 1078 SRIALGPKRGPSVTEKEVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELIS 1137  
DB 1081 SRIALGPKRGPAVTEKEVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPEVDHL 1140  
QY 1138 GEALDPLFMDPDLAYGTYTSGCGHVMHVCWKYFEAVQLSSOORIHVDLPDLESGEYLC 1197  
DB 1141 GETLDPLFMDPDLAHGTYTSGCGHVMHVCWKYFEAVQLSSOORIHVDLPDLESGEYLC 1200  
QY 1198 PLCKSLCNTVPIPIPLQPOKINSENADALAQLLTLARWITQVLARISGVNIRHAKGENP- 1256  
DB 1201 PLCKSLCNTVPIPIPLQPOKINSENADALAQLLTLARWITQVLARISGVNIRHAKGEAPA 1260  
QY 1257 IPIFFNQMGDSITLFPHSILSPGVSSIKYSNISIKEMVILFATTIYRIGLKVPPDERDR 1316  
DB 1261 VPVLFNQMGDSITLFPHSILSPGVSSIKYSNISIKEMVILFATTIYRIGLKVPPDLDPR 1320  
QY 1317 VPMLTSTCAFTIQAIENTLGDGKPLFGALQNRQNGLKALMQFAVAQRATCPQVLIQK 1376  
DB 1321 VPMMTWSTCAFTIQAIENTLGDGKPLFGALQNRQNGLKALMQFAVAQRATCPQVLIHK 1380  
QY 1377 HLVRLLSVLPNPKSDETPCLSIDLPHVLVGVAFAPPSLYWMDPDVLOPSSVSSYNHL 1436  
DB 1381 HLARLLSVLPNPKSENFTGILLSVDLPHVLVGVAFAPPSLYWMDPDVLOPSSVSSYNHL 1440  
QY 1437 YLFHLITMAHMLQILITVDT-----GLPLAQVQDSBEAHASAFPFAEISQYTSGSTGCDI 1492  
DB 1441 YLFHLITMAHMLQILITVDT-----GLPLAQVQDSBEAHASAFPFAEISQYTSGSTGCDI 1500  
QY 1493 PGWYLVSLKNGITPYLRCAALFFHYLLGVTPPEELHTNSAEYSALCSYLSLPTNLFL 1552

DB 1501 PGWYLVSLKNGITPYLRCAALFFHYLLGVAPPRELPAASAEGEFPALCSYLSLPTNLFL 1560  
QY 1553 LFQEWYDVTIRPLQWCADPALLNCLKQKNTVTVPRKNSLIELPDDYSCLLNQASHPR 1612  
DB 1561 LFQEWYDVTIRPLQWCGDPALLKSLKQKSAVVYPRKNSLIELPDDYSCLLNQASHPR 1620  
QY 1613 CPVSADDERKHPVLCLFCGAILCSQNICCOEIVNVEEVGACIFHALHCGAGVCIFFLKIRE 1672  
DB 1621 CPVSADDERKHPVLCLFCGAILCSQNICCOEIVNVEEVGACIFHALHCGAGVCIFFLKIRE 1680  
QY 1673 CRVVLVEGKARGCAYPAPYLDVEYGETDPGLKRGNEPLHLSRERYRKLHLVWQOHCIEETA 1732  
DB 1681 CRVVLVEGKARGCAYPAPYLDVEYGETDPGLKRGNEPLHLSRERYRKLHLVWQOHCIEETA 1740  
QY 1733 RSQETNQMLFGFNWQLL 1749  
DB 1741 RSQETNQMLFGFNWQLL 1757

## RESULT 2

T25604

Hypothetical protein C328.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T25604

R:Gatung, S.

submitted to the EMBL Data Library, February 1997

A:Description: The sequence of C. elegans cosmid C3288.

A:Reference number: Z20056

A:Accession: T25604

A&gt;Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1927 &lt;GAT&gt;

A:Cross-references: UNIPROT:P91133; EMBL:U88308; PIDN:AAB42328.1; GSPDB:GN00019; CESP:C3

A:Experimental source: strain Bristol N2; clone C3288

C:Genetics:

A:Gene: CESP:C328.11

A:Map position: 1

A:Introns: 56/3; 157/3; 255/1; 469/1; 578/3; 832/3; 907/2; 1165/2; 1218/2; 1275/2; 1468/

Query Match 21.1%; Score 1945.5; DB 2; Length 1927;

Best Local Similarity 27.6%; Pred. No. 2.5e-111;

Matches 560; Conservative 311; Mismatches 663; Indels 493; Gaps 66;

QY 97 QLCGRVFSGETTYSRCDCAIDPTCVLCMDQFQDSVHKHRYKMTSTGGGFCDCGDTGA 156

DB 14 QICGHVFKNGELTYTCLDCATDGTCTVMCLQCPEVSIHKSHKYMHSMSGSGYCDGDA 73

QY 157 WKTGPPCVNHEPG-----RAGTIKENSRCPLNEEVIVQARKIPPSVIKYVEMTWIER 209

DB 74 WTEGYACANHEKODDEAAVLAPELKK--RCSQLVEIILQ---FSLSMITHKDDLLKPEI 128

QY 210 EKELPELQIREKNERYCYVLFNDEHSHSYDHYVYSIQRALDCELAEQAQLHTTAIDKEGR 269

DB 129 FEKMEPE--VTNEAQOYLTVLYNDEHTYVESIVKLELYIHTCKQOAMLVATVDREGS 186

QY 270 AVKAGAAACQBAKEDIKSHS-----ENVSOH--PLHVEVLHSEIMAHQKPAKRLGWS 320

DB 187 AVKLGSKADCTKAKDDVQRKTARDPTSTRSSNNHNLPLSVKVMDDTTLFALQNFSLIATW 246

QY 321 MNKIMSYSDFRQIFCQACL----- 340

DB 247 LNTQMDVFPPLREIVGEIILLSSKFALKKNYTRMKSEDRQLVAGIIRNVNVLPPDEEEL 306

QY 341 -----REEP-----DSENPCLI 352

DB 307 PALDGRMDVEMDDDDIDIGEALQMEIDADDEBEITAAALAGVSEHQSSPGPSRDSSTFTML 366

QY 353 SRLMLWAKLYGAKKILHELFPSSFFMMEYKGLFAMEFVKYKQLQKEYISDDHDSRI 412

DB 367 ENILLQDTQMWKAGRSILHMLMRTVPMIYDQKVPKAFKMLHYNEIYEDFTKODHEMDV 426

413 SITALSVOMPTVPTLARHLIEQNVISVITETILEVLPEYLDNRN-----KFNFGQYS-Q 466  
427 SVVGLSVQMTVPSLARKVAEQAFSVISKALRQDTDFVKTYNDGKIARPDFTSRSP 486  
467 DKLGRVYAVICULKYILISKP--TIWTERLMQFLEGFRSLKILTCMQGMBEIRRQVGQ 524  
487 PELRRSLHITRDWAYILNAVPSSESDNRELIDGFGVQFADFLFLQHLQGMDEVKQAVE 546  
525 HIEVDPDWEAAIAIQMLKNILMFORWCACDEL---LLVAYKECHKAVMRCSSTFIS 591  
547 HQWSESEWETAFFILLRLKDAISMIIGWAFNEEVNRLMIM---CLELNRMPVYTKS 603  
582 S-----KTVQSCGSHLETSKYSRVSDELVSILPLSRTLGL-----HV 620  
604 BEDTYELTVINGESCRIHDFVDKSSTSVHQVVRRIIAGLFSASNTYFGLNRNSNST 663  
621 RLRLGAVARLHEFVFEFQVVLVEYPLRLCLVLVAQVVAENMRNGLSLISQVYYQD 680  
664 SLNQ-----ERIKELINCKO--ETNLAYELSLVLVLCASNATLWRNGFSLINQIHNYS 717  
681 VKCREMYKDIIMLOIGASLMDPNKFLLLVLORYEL-----ARAFNTISTKQD 730  
718 PLCRNEMFDRDLVMQVGAALTSTKFIPIHLHLRFLFKWATSEFPQDKANEKPAKPESE 777  
731 DLIKQYNTLIEEMLOVLIVYGERVYVGVNVTKEVTVREIILHLACIEPMHPSATAKNL 790  
778 DLSKTLVIAEETPCILLICERYTYGVGKTRPMQDKREVJHILCTGSHTFSHIQKM 837  
791 PENENNE--TGLENVINKVATFKKPGVSGHGVYELKDESLKDFMYFYHYSKTOHSAEEM 849  
838 SHDINSKRSLHEAVNLVADFRPLATTAGQFCKESSLPTTYPFPWYHYSKDSQSAEQS 897  
850 QKK--RRQENKDBALPPPPPEFCAPSKVINLLNCDIMMYILRTVFERAIDTSDNLWTE 908  
898 QARVRAKMEKSVACAPPILPDFQTFPERIPEILTTNVLHVLLIIRTA--RRSRFSD 956  
909 GMLQMAFHIALGLLEEKQOLQVAPEEVTFDYHKAISRLGSSAMNIQMLEKLGIPOL 968  
957 RLPHKTYLIGITALNEE-----EKNPSFGTORA-----EESIGLLSLLEGLVCKP-- 1002  
969 EQQKDMITWLOMFDVTKRLREKSLIVATTSSSES--IKNDRIETHDEKAEKRAKAA 1026  
1003 --ESSICPLLEV--TVEKTRK--LIKARAGVPEAPAPENKPAQSEBEIKAKRARA 1055  
1027 RLHRQKIMQMSALQKNFIETHKLMYD---NTSEMPG-----KED-----SIMEERS 1070  
1056 EM-RQAKAMKSNMQSKFM---KKIEDEEKQDESQTPSEKSETVVKQDYDNKHPFDEDV 1111  
1071 TPAVSDYSRIALGP-KRGPSTVEKEVLTCILQEBQEV-KIENNAMVLSACVQKSTALQ 1128  
1112 VKQVGHDFPVCIGANKWHAELVKPRTLTCILQEDBIIAPQOGKPVCAAFIQOSQLFT- 1170  
1129 HRGKPIEL---SG--EALDPLMDPLAYGTYTSCGHVMAVCHQKYFEAVOLSSQOR 1182  
1171 HKNKGELMTASSGTTSTRDLLTAPATLOYGVNDVSTCSHSMHYECYRSIAENRSRESLR 1230  
1183 IHV-----DLFDESGEYLCPLCKSLCNVPIIP---LOPOK-----INSENADAL 1226  
1231 ARQVGQSHKWVDTEGEGEQCLCKELSNAAIPVLPAYQLTQWNGFSTVSGAKENFDT- 1289  
1227 AQLLTLIARVITQVLARI-----SGYNIRHAKGNPI-----PIFFNOGMG 1266  
1290 -----WVARVKNRLEMPLESVSKGHSKRKRSERSLIDLEKSKDPDTANTSAG 1341  
1267 DSTLEFHSILSRV-----ESSIKYSNISKEMI----- 1295  
1342 -----VLQFQAMGMSATHMPASABSOQLMTTSPSQDDVEFYNELAAMPVQDVNNT 1393  
1296 ---LPATTIYRIG--LKVPDPDERDPRVM----- 1319  
1394 TSPAATPETIPATGSSSRIPESQESGKPLSSOIQHVLXSLRPPFALLNSNRICSSSPE 1453  
1320 -----LTWSTCAFTIQAIENTLIGDGE 1340

1454 GFEEPIKDLGKNMKFRKGNELKTNFIEKHLKGYVISTVWQSTAHVARALSSYLHYDK 1513  
1341 KPLFGALQNRQNGKAL-----MOPAVAQARITCPQVLIQKHLVR--LLSVVL--- 1386  
1514 KPLFGALNRQRDCLSAMARLCSLSHNQFLL-----HAVSDMLRVFLCEP 1560  
1387 PNKSEDTFCLLSIDLPHVLVGNVLAAPSLYDDPDVLOPSSV-SSSYNHLIYFLHITWA 1445  
1561 PRPKLAQTP-----GSLPLSAPSTSSFTPA---PAQIPHSGTNFAFLVLFNPA 1606  
1446 -----HMLQILLTVDTGLPLA-----OVQDSESEAHGASSP 1476  
1607 GPRKNVNLILQIDI-----LSLAIATHSADNDVNMEEQESQESMEVDVPAQAIRKL 1661  
1477 PAEISQYTSQSGCDIPGWYLVNKLNGITPYLRCAALFFHYLLGVTPPEEL---HTNSA 1533  
1662 YALCHPFDGDLRRVDI---LWRKMBEGAQSLRLPIALLYHFTLVPPEALKDPSINGS 1717  
1534 EGEYSALCSYLSIP-----TNLPLFQBYWDTVRPLLQRWCADPALLNCKQK 1581  
1718 E-----PLFRYLGLPHKIEQISGSMLEKLFTH---WSSSIP-----SDQAL---R 1757  
1582 NTWVRYPRKRNLIELPDDYSCLLNQASHFRCPRSADDERKH--PVLCLFCGAILCSQMI 1639  
1758 QDLVQVQVRPNLLVELPEKYSQILNQVATFKCPTPIESTSNVPTLCLVCGTILCSQAY 1817  
1640 CQOEIVNGEVEGACIFHALHCGAGVCIFLKIIECRVVLVEGKARGCAYPAPYLDREYTD 1699  
1818 CCKIINKQSYGACRYHMSQCSGSGVGMFLVRDCSLVMTTRKRGCFRPAPYVDFEGEVD 1877  
1700 PGLKGNPLHLRERYRKLHLVWQHCIIIEETARQETNQLMFGFNW 1746  
1878 QGFRRGNPLHNPVLYQKLKSLWLOQGITVEEVNNEIDFRNVQYDW 1924

## RESULT 3

Hypothetical protein SPBC19C7.02 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T39808

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gaillardin, C.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z21881

A:Accession: T39808

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1958 &lt;LYN&gt;

A:Cross-references: UNIPROT:O60152; EMBL:AL023859; PIDN:CAA19570.1; GSPDB:GN00067; SPDB

A:Experimental source: strain 972h-, cosmid c19C7

C:Genetics:

A:Gene: SPDB:SPBC19C7.02

A:Map position: 2

Query Match 11.3%; Score 1038; DB 2; Length 1958;

Best Local Similarity 23.3%; Pred. No. 2.4e-55;

Matches 480; Conservative 312; Mismatches 763; Indels 502; Gaps 88;

60 EKQESVQMSIPTPLEWY-----LFGEDDPICL-----EKLKHSAGFQLC 99

37 ESAKESLLNEVFSALLGYDHTLWNTLLPERPTIDASFLLRRAQGHSEGEYRHRGTCSKC 96

100 GRVFKSGTTCRCDAIDPTCVLQMDQFQDSVKNHRYKMTSTG--GGFCDCGDTFAWK 158

97 GHIFRAGEVYFRCKTCSVDNSALCVKFRATSHKDHETSFTVSAGSGGCCCGGANAAM 156

159 TGPFCVNHPEGRAGTIKENSRCPLNEEVIVQARKIFPSPVIKYVENTIMEEB--KELPPE 216

157 GDVSKIHSHEDATISNDMIDEIEKLENSIQTTIDCVLDVLDVFCSPENLKQMP 216

217 LQI--REKNER-----YCVLFNDRHSHYDVHYISLQALDCELAE---A 256

Db 217 ESILQDEKTSRLSENKYGIDDDSCNNYSILVWLNDEKHSPKFQYEQITTA--ELPNNVFG 274  
QY 257 QLTHTAIDKEGRANVAKAGAYACQAKEDIK---SHSENVSOHPHLHVEVLHSEIMAHQKF 313  
Db 275 KXVANIINDIGR-----ACIVTETNIKELLKIGOKLAQINLAVSIRSRMDIFREES 325  
QY 314 ALRLGSMNKMYS-----SDFRQIFCOACLRL----- 341  
Db 326 CAVILLEWADIAGSSICGRNYPSSVICKELVRPNWCGLNHSLDPLRLSLRLAUPAIVA 385  
QY 342 -BEPD-----SRNPLISRLM----- 356  
Db 386 IDSPDIFLNEDHNSGSDTSSHMLTDESSIHRSRWYPSNSLPDVLVSASVRVDFYFF 445  
QY 357 LWDAKLYKARKILHLEIFSSPFMEMEYKLFAMEPVKYKIQLOKEYISDDHRSISITA 416  
Db 446 LYDLKLWLSRLYKLOELYLGYFITOPGFKBEIMGARIALSYRRLAELFLLLDREPEHSVIP 505  
QY 417 LSVQMETVPTLARHLEEQNVISVITETLVL-----PEYLD-----RNKNFQGY 464  
Db 506 FSWQIPTVADVAKGLVTEYDFLTINATLYTFTYKKLNTPNYVDQHAMIRDSAAFH-- 563  
QY 465 SQDKLGRVAVICDLKYIILISKPTIWTBRLR--MOFLEGRSFLKILTCMQGMEIRROV 522  
Db 564 -----SRRYIHIFHHIQFMLSIPCV-AEIVREDLKFKQYADFNL---FQGMCPVTRAV 614  
QY 523 GQHIEVDPP-WERAAIAIQML-----KNILLMPQEW-----CACDELLIVAYKECHKAV 571  
Db 615 SQHVEWENDSMYVNLVNSLQVAKLCHRGVGNVFMELNTKLANAINYLISLILYPKARNES 674  
QY 572 MRCSTSFSSKTVQSCGHSLETKSYRVSDELSVTHLPLSRTLACGLHVLRLGAVSRL 631  
Db 675 W-TNTESLTGTITVDER-GNS-KLIBYDIALQPVSFHHPLHMLVYL-----L 719  
QY 632 HFEVSPEDF-----QVEVL--VEYPLRLCLVLAQVVAEMWRNGLSLISQVFFYQDVKCR 684  
Db 720 SPVVERDNYKLLWTQLDLLAVTDHPLRVCAWLSQMBKAWIRNGTTLRDQAHYRNLSFH 779  
QY 685 BEMYDKDIIMQIGASIMDPNKFLLVLORYELAEAF--NKYISTQDOLIKQYNTLIEE 742  
Db 780 EYTFDLVLLQLLTGYDPDAILPSFISRFQLEDQMGYRFPVPHKHYD-VSQVTIMBEE 838  
QY 743 MLQVLIYIGERVVPGVGNVKEEVTMBELIHLCLTEPMHPSHSAKNLPENNENNTGLEN 802  
Db 839 FLLLLSIIVCNTAVLDHMDITRR--TEYGLAHILCFRPLPYSITRTCEHLLHKKQFES 896  
QY 803 VINKVATFKK-PGVSGHGVVELKDESLKDFNMFFYHYSKTQHSKASHMOKKR-RKOENKD 860  
Db 897 TLKKVATFRNAEINDSGSFTLDEYFDYVDPFNIHYSRNQREAEANILRRRYSKQHSKH 956  
QY 861 EALPPPPPPFCFAPFSK--VINLNCDDIMMYIL-RTVFERAIDTDSNLMTEGMLQWAFHI 917  
Db 957 --LESVYVEBYPHLSHNITIPILQSDSFVGLWHTIVVAYIYPYDQKLEGLVNTALHA 1014  
QY 918 LALGLLEEKOQLOKAPEEVEVTFPYHKASRLGSSAMNIQMLLEKLGKIPQLEQKD---- 973  
Db 1015 CLLVLMSK-----GSEPIFSKKICE--NRFPVVEGLQBYCNS 1050  
QY 974 ---MTTWILQM----FDTVKRLREKSKCLIV-----ATTSGSESIKNDE 1009  
Db 1051 PDVTLFSVLQMKQNRHNFVYVK--EKISLIMKILKSEVPLLYEPVVAETLSISSSKIYQ 1107  
QY 1010 ITHDKBAERK-RKAERARLHRQIKMAQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEE 1068  
Db 1108 SLSDAEQEQHLAKVRMAKERQARIMEQPRMQNKFLENHALFEASDCEN----DEADEF 1163  
QY 1069 ESTPAUSDYRIALGPKRGFSVTEKEVLTICILQEREQEVKIENNAMVLSACTQKSTALT- 1127  
Db 1164 SVTSSVS--TKJFLDP-----PIDTCLLQCEELKDKRPYGTTLVF---VLRSSVRL 1209  
QY 1128 -----OHRGKPIELSG--EALDPL----- 1144  
Db 1210 FPADDANYSEVLIDIPDLSLDBHEIQERFPFLAGKKKVKVLDSTEAYDYDNYEYKKGKLNHQ 1269

QY 1145 -----FMDPDLAYGTYTSCGHVMHACVCKYFEAVOLSSQ-----QRIHVDLFLDESSE 1194  
Db 1270 LKDSFNGPPDQDGLHATCGHFHMDCFKNHITATVTLATRANPYRNHP--HNLSMKE 1327  
QY 1195 YLCPICKSLCMTVPIIPILOPKINSENADALAQLLTLARWI-----QTVL-- 1240  
Db 1328 FLCPLCKALCNTIFPILWRPKBEINFQAGVLT--APLKNMLVSKTSFNFKNOLNQQLLDI 1385  
QY 1241 -----ARISGYNIRHAKGENPIPIFFNQMGDSTLEFHSILSFGVLESSIKYNSIKEMV- 1294  
Db 1386 ETPSEHTQSYNL-----NLDDV-LQHTLRDLSKDIYT-LNTGADNS--SDNVEENAD 1434  
QY 1295 ILFATTIYR-----IGLKVPDPDERPRVPMLTWSTCAF-----TIOAIENLLGDE 1339  
Db 1435 NLFQSSVLVDHVFHSVNVNEVPADER-----LAISDDIFELYRLRDDVDIDLNSSLYSDD 1488  
QY 1340 GKPLFGALQNRQHNGLKALMQFAVAQ-----RITCPOVLI-----QKHLVRL 1382  
Db 1489 PIPVNGKL---HNVVK-LFSVSLCQVEASTRGHIKCSSIPADIWVHNLGKNQVFLRL 1543  
QY 1383 SVVLENIKSEDTPCLLSIDLPHVLVCAVLAP-----PSLYWDDPV--- 1422  
Db 1544 S---ESIKTYTLLC--AHDQKRIGGSIQEFEFISFCOOKRIFGRLLPSL--DSPVTKSI 1596  
QY 1423 -----DLQPSVS-----SSYNHL-YLPHLITMAHMLQILLTVDLTGLPL 1460  
Db 1597 TDDRVEPLLVKDTFREFABASVSGLLSCDESFPHYLTQLYYTIADIVRNMLWILLSQRNSL-L 1655  
QY 1461 AQOVDSDEA-----HSASFFABISQYTSGSIGCDI-----PG--WYLWVS 1500  
Db 1656 KCMSEVFEAPDYEQKGFELHVIQIWKSLRVGDAGLINFDCCTBDDLMNPHLLFTLYKL 1715  
QY 1501 LKNGITPYLRCAALFFHYLLGVTTPPEELHTNSAESEYSLCSYLSLPTNLFL--LFQBY 1557  
Db 1716 LERFSLIFLRKCALLMYCRYGVSPETQPNLNFQNSLSRLQTKMHIPGVIELSNHLCLTA 1775  
QY 1558 WDTVRPLQRWCADPALLNCLKQKNTVVRYPKR---NSLTELDPDYSCILNQASHFRK 1613  
Db 1776 STESWLSLHKWC-----NFFTETGPLCDPPRAYYPGIYELVSLPYELDKVPELLARAC 1829  
QY 1614 PRSADDERKHPVLCLFCGAILCSQNICQEI VNGEVEGACIFALHCGAGVCFIKIRBC 1673  
Db 1830 SKCL-TEPMEPAICLFCGKLLCFQSHCCS--FNG--IGECNLHMQOCASDIDGIFLIVKRC 1884  
QY 1674 RVVLVEGKARGCAYAPYLDYEGETDPLKRGKNPLHLSERY-RKLHLVWQOHCILEETA 1732  
Db 1885 AILYLNPVY-GSFSVAPFLDAYGETDLGLRRGRSQYLSQRYDETVRTVWMLNGSIPSYIA 1943  
QY 1733 RSQETNQMLPGFNWQLL 1749  
Db 1944 ROLDANPDFTGG--WETL 1958

RESULT 4

T37711  
probable n-end-recognizing protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T37711  
R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, June 1997  
A:Reference number: Z21738  
A/Accession: T37711  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2052 <MUR>  
A:Cross-references: UNIPROT:O13731; EMBL:Z97208; PIDN:CAB10108.1; GSPDB:GN00066; SPDB:S1  
A:Experimental source: strain 972h-; cosmid c15A10  
C:Genetics:  
A:Gene: SPDB:SPAC15A10.11  
A:Map position: 1

Query Match			9.9%	Score 914;	DB 2;	Length 2052;
Best Local Similarity			20.4%	Pred. No. 1.2e-47;		
Matches 447;			Conservative 326;	Mismatches 734;	Indels 682;	Gaps 89;
QY	62	QBSVQMSIPTLEW-YLFGEDEP-----DICKELKHS-----GAFOLCGRVF 103	DB	1005	-EQILHVDIODIIRTRLAQQLCFGLPLAYSALLSTISSNLVSELSPOKIREVTSYKAPDG 1063	
DB	43	QEVPLSLNVDNWKYFLKEPGAITSDPFLSRLOHSEPECAQELQDKRSGSKVCGHVF 102	QY	815	VSGHGVYELKDESLKDFNMYFYHYKTKQHSKAEHMQKRRKQEN--KDEALPPPPPEFC 872	
QY	104	KSGETTYSRCDCAIDPTCVLCMDCPQDSVHNKRYKMTSTG--GGFCDCGDTBAWTKGPF 162	DB	1064	LHDPGVYSLKDEYDVLVDPYFYHYNKNREESOTILKRLAKKNVSAESIIEPKIRPL 1123	
DB	103	RAGEVIYRCNCGLDNTECVLCAPCFRATNHEGHETHVSISTSYSGICDCGDPRAWNVDLN 162	QY	873	-----PAFSKVINLLNCDIMMYILRTVFERAIDTSDNLWTEGLQM-AFHILALG--LLE 924	
QY	163	CVNHEPGRAGTIKENSRCPLNEEVIQARKIPPSVIKYVEMTIW----- 207	DB	1124	EKQGHDIPTAAVNAFTSLIIF-----RAIE-----YALVQAESFGSSDIGNTILG 1169	
DB	163	CKTH-----NVPPDEBKQPEVI-----PLEQHSIRTTIHILLDFILDFVFCSPV 209	QY	925	EKQOL-----QKAPBEEVTFDFYHKA-----SRLGSSAMNIQMLEKLGKIPQL 968	
QY	208	-----BEEKELPPELQIREKNERY-----YCVLFNDEHSHVDHIYS 244	DB	1170	DALQCLISMKIHFPKSNDFCSRCAERYPTDSSIMRFBGGSAYCLAELCFAILKSPKY 1229	
DB	210	NLXAQSTVSGIILADEAS-----RLSAKYGIVADRPCNVRVLMWDEVHTFDVAVGS 262	QY	969	EKGOMITWI---LQWFD--TVKRLREKSLIVATTSGSESIKNDITTHDKKAEARRKA 1023	
QY	245	LORALDCELABEAQLHTT-AIDKEGRRAVKAGA----- 275	DB	1230	KOVHVKNVAVLAGLQKNDPAYSNNLEATHPELSTTSS-----TSDSNEIEKTQEK 1280	
DB	263	VLEALDSSNTAFGLVAQRVDSIGRPVATSAVSHEAIRIANAIKENLAVNVRTARDPP 322	QY	1024	EAARLHQ--KIMAQMSALQKFIETHKLMYDNTSEMPGK---BDSIMEBESTPAVSDYSR 1079	
QY	276	-----YAA-----COE----- 281	DB	1281	KRLALEKQKKIMQOFRDQOASFLA-----QNTDFDIBGDTEDVTTEPEBEVKYHEH 1334	
DB	323	REDICGILFEWFDLLESHVCYFADYLQIIVCDEILKNWSPGLEKPAKPEVNFNPLPEI 382	QY	1080	IALGPKRGPSTVEKEVLTICILCOEQQEVK-----IENNAMVLACVQKSTALTQHRG 1131	
QY	282	-----AKEDIKSHSENVS-----QHPL----- 298	DB	1335	I-----RG-----NCLCOEBCNDQAPYGVIGIIOGSSLLRKTVDHSEILDEIYS 1380	
DB	383	VNDDESDDIYAEELLDVIANLODETGVTRIANLGGDEPDEADMTDPTTAGDFHPLDD 442	QY	1132	KPIELS-----GEALDPLFMDPD-----LAY-----GTYTSCGHVMAVCWQKY 1171	
QY	299	-----HVEVLHSEIMAHQK-----ALRLGSMWNKIMSYS----- 328	DB	1381	VPPNLDRESHSRPPKKYDVTVFNRSKORLLSAYPGNNIRGVFVSGCGHMLHLCGFKY 1440	
DB	443	NDVNDLLDPETEREDIDDLTDEVMEETEENAEADYGVNVRNTRQDDVDQDISMETESQNE 502	QY	1172	FEAVQLSSOORIHVDLFDLESG--EY-----LCPLCKSLCNTVPIPIIPLQPKIN 1219	
QY	329	-----SDPQIIP----- 335	DB	1441	YVARSMYRN-----DVTAGLSEYYKYSTAKPFMCPCLCRSLSNVLLP-MPQIPKM-- 1489	
DB	503	TDESQNTENVYNPQTHPTVPIPTATQDVVTIRPEFNSQLNLRQIINARRPRPAAV 562	QY	1220	SENADALAQLLTLARWIOVTLARISGYNIRHAKENPIPIFPNQGCGDSTLEFHSI---- 1275	
QY	336	COACLREE-----PDS-----ENPCLISRL--MLMDAKLYKGARKILHELIFSSPF 379	DB	1490	CLNIDTLPFRSMNGWLEEI--GTMSSSFEYQ-----LVRSSLSDTKOTFRSCFLRP 1540	
DB	563	CQVSLREDYKSPHPPTPPSSYSVFESPSSILRDYFLFDLKFWKRLGLRLSKLYVVPFN 622	QY	1276	-----LSFGVESSIKYSNS--IKEMVILFATTIVIRIGLKVPPDERDPRV-PM 1319	
QY	380	MEMBYKCLFAMEVVKYKQLOKEIYISDDHDSRSTISALSQVMTPTVLARHLIEEQNVIS 439	DB	1541	WTNSKIISAMLAELKIADGALIDQSNRNDVSDLYDRYCETT-KLANKLVKGSTFTVNSPH 1599	
DB	623	RNLFPKRLGIRFVIHYSRLATAPLADREPDSHVMFLSVQPTTISLAEAVKVDPLT 682	QY	1320	LTMSTCAPTIOAIENLLGDEKP-LFGALQNRQHNGKAL-MQFAVAQRITCPQVLIQKH 1377	
QY	440	VITETLEVL-----PEYLDNRKNFPQGYSD-----KLGRTVAVICDLK 480	DB	1600	DLINSLAYTVSSLE--VSQRCSPKQSGATRSVWFNELGFLTSLFPLTSDT----- 1648	
DB	683	NLNATTLSSLTQSNRPSTL-----FSSDIETPTIQLNRQVLRKTRTYNLPFLDLG 732	QY	1378	LYRLLSVVLVNTKSEDPCLLSIDLHVLVGVAVLAPPSLY-----WDDPVDLQPSV 1429	
QY	481	YIL-----TSKPTIWTBLRMOPLFGRSPFLKILTCMOGMEIRBQVQHIEVDPDHEAI 536	DB	1649	---VLKVCVDQIISKDQOALLMESQKLJCVCKI FYRHSQKSLMRNGRMSHDQIQPPLL 1705	
DB	733	YLLQHPQVKKLVDDTRYVHQYIDLLRVF-----QGVIPQORAILSHVQ---WDFPH 781	QY	1430	SSSY-----NHLVLFHLITMAHMLQILLTVDTGLPLAQVQED---SEE 1469	
QY	537	AIQOMQLNILLMPQEWACACDEEL-----LLVAYKECHKAVMRCSTSPSSKTVVQSCGHS 592	DB	1706	SNTFDDFVKISSMLMIFGQDNILTYKLVFYLSBKTIISMIKKVVADSVVPDLTINS 1765	
DB	782	G-----KNILFVMQVRVAMLSNVTSSCFTQAPYERLFVAI-KCIITSITHPKLDI---AES 832	QY	1470	AHSASSFABISQYTSIGSGICDIPGWYLVWSLNGI-----TPYL 1509	
QY	593	LETKS-----YRVSEDLVSIHLPLSRTLAGLHVRSLRGAVSRHLHFVS 636	DB	1766	QOSKQOY-----ILC--KNVLLWCGSSNNIIEILDDBSNLLRLMSLVKEYSLPFL 1813	
DB	833	LEPLSCIPSSLTNTQPLVPPSVSRDPIFSYHP-----LHWMLSNLFSYCRVDASSH 885	QY	1510	RCALFPHYLLGVTPP-BELHTNSAGEYSALCSYLSLPTNLFLPQFW-----DT 1560	
QY	637	FEDPQVEVLVEPLRCLVLAQVVAEMWRNGLSISQVFPYQVCKREEMYDKOIMLQ 696	DB	1814	RRVALVLYCMFDISLEFNEFBSNNEDDSELERLSKLIKVP--LQELYSQMSDENNQ 1868	
DB	886	WDKOTLALLDHLPRVCVLLAQIDCNLWTRNGRSILLTDAFYRQLNIEVSYDKOILAIO 945	QY	1561	VRPLQRWCADPALLNCLKQKNT-----VVYVRPKNSLIE-LPDDYSCL 1604	
QY	697	IGASLMDPNKFLILLQRYELAB-AFNKTIISTKQDLIKQYNTLIEBMLQVLIYIGERY 755	DB	1869	ILEIAGWCEHLA-----QNTWGDSTISLEYPGIYELVKUPLHLENLIDSMQSVCCM 1921	
DB	946	TILMFVDPMLVNAVQREFTDWLNTYNEHPNVDTERIPAMUCMKMLLELLIALITER- 1004	QY	1605	LNQASHFRCPRSADDERKHPVLCFLCGAILC---SONICQCBIVANGEYVACIFHALHCG 1661	
QY	756	VFCGVNVTKEVTMBRIIHLCTIEPMPHSAIAKNLPENNETGLENVINKVATFKCP-G 814	DB	1922	CHKTPL-----PAICMLCGSVICFNAQNTVSSRLTGE-----CNKHAATCT 1965	
			QY	1662	AGVCIFLKIRECRVVLVEGKAR-GCAYPAPYLDVEYGETDGLKRGNPLHLSRERY-KLH 1719	

Db 1966 GSVGIFPITKACGILLLLDSISNTGTTIMPTPYLDIHGETDLQLRGCGPQLNCKRYDFVVR 2025

Qy 1720 LWQOHCIIIEETARSOETNOMLFGFNQL 1748

Db 2026 EQWLQTVLQKMARHMDTMEQ---NWRM 2051

RESULT 5

S12332

ubiquitin-protein ligase (SC 6.3.2.19) - yeast (Saccharomyces cerevisiae) (strain S288C)

N:Alternate names: protein G7168; protein YGR184C

C:Species: Saccharomyces cerevisiae

C:Date: 16-Sep-1992 #sequence revision 16-Sep-1992 #text\_change 09-Jul-2004

C:Accession: S12332; S64502; S64498

R:Bartel, B.; Wuening, I.; Varshavsky, A.

EMBO J. 9, 3179-3189, 1990

A:Title: The recognition component of the N-end rule pathway.

A:Reference number: S12332; MUID:91006011; PMID:2209542

A:Accession: S12332

A:Molecule type: DNA

A:Residues: 1-1950 <BAR>

A:Cross-references: UNIPROT:P19812; EMBL:X53747; NID:G4743; PIDN:CAA37779.1; PID:G4744

A:Experimental source: strain S288C

R:Arriyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64499

A:Accession: S64502

A:Molecule type: DNA

A:Residues: 1-1950 <ARR>

A:Cross-references: EMBL:Z72969; NID:G132325; PIDN:CAA97210.1; PID:G132326; MIPS:YGR18

A:Experimental source: strain S288C

R:Hebling, U.; Hofmann, B.; Delius, H.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64003

A:Accession: S64498

A:Molecule type: DNA

A:Residues: 1615-1950 <HEB>

A:Cross-references: EMBL:Z72969; MIPS:YGR184C

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:UBR1; PTR1

A:Cross-references: SGD:S0003416; MIPS:YGR184C

A:Map position: 7R

C:Superfamily: ubiquitin-protein ligase

C:Keywords: ligase; protein degradation

Query Match 7.5%; Score 689.5; DB 2; Length 1950;

Best Local Similarity 19.8%; Pred. No. 7.5e-34;

Matches 409; Conservative 323; Mismatches 732; Indels 601; Gaps 86;

Qy 7 GGTREMEISALPQTPQRLASWDDQVDFYAF-LHHLAQLVPEIYPAE-----MDPDL- 59

Db 33 GPTERADMSRALKEFIYR-----YLFVINSGENLPTLFNAHPKQKLSNPFLT 81

Qy 60 ---EKEESVQMSIFPLEWYLFGE-DPDCLEKHLXSGAFQCGRVFKSGTYYSCRD 115

Db 82 VFPSLSDAVIDDKITSQQTIPYKIDESRIGDVHHTG--RNCGRKFKIGEPLKYCHEC 139

Qy 116 AIDPTCVLCMDCPQDSVHKHRY--KMHSTGTGGFCDCGTEAWKTPFC-----VN 165

Db 140 GCDDTCVLCHCFNPKDHNHVVCTDICTFTSGICDCGDEEAWNSPLHCKAEEQNDIS 199

Qy 166 HEPRAGTIKENSRCPLNEV-IVQARKIPPSVIVVEMTWIWEBEKELPPELQ----- 218

Db 200 EDPATNADIKEED--VWNSVNIALVELVAEFDYFID--VFQNIETPLPTIQDKITIK 255

Qy 219 IREKN-----ERYVCVL 230

Db 256 LEWMTQGRMYERAQFLNDKYENDYMGDTTAKTSPNSPEASPLAKIDPENTVII 315

Qy 231 FNDHEHSYDHVIYSLQALDCELAELTAIDKEGRRAVKA-----GAYAAQCE- 281

Db 316 YNDEYHNYSQATTALRGVP-DNVHIDLTSRIDGEGRAMLKCSQDLSVLGGFFAVQTN 374

Qy 282 -AKEDIKSHSENVSOHPLHVEVL---HSEIMAHOKFALRLGSMWNKIMSYSSDF----- 331

Db 375 GLSATLTSWSHYLHOETCKYIILWTHCLNIPNSSFOITFRNMCKTL---CSEYLNATEC 432

Qy 332 -----RQIFCOACLBEP-----DSN 348

Db 433 RDMTPVVEKYFSNKFDPNDPYRIDLSILADGNCQIPLGHHKILPESSTHSLSPINDVET 492

Qy 349 P-----CLISRLMLDAKLYGARKILHELIFFSPFMEMEYKGLPAMEFYKYQLOK 401

Db 493 PSTRYSNTRLQHILYFDNRWKRLKDIQNVIIPTLASSNLYKGFICQOVVVEIFNHNTR 552

Qy 402 B--YISDDHDSISITALS---VQMTVPTLARHLIEBQNVLSIVTETLLEPEYL--- 453

Db 553 SVAYM-----DREPQUTALRECVVQIPTCTNAKNIFENQSFIDIVW-SIIDIPKFECKVE 607

Qy 454 -----DRNNKFNF-QGYSQDKLGRVAVICDLKYLISK---PTIWTERRLMQPLEGPR 503

Db 608 GGVLIWQVRKSNLTKSYISFKQGLYTV-----ETLLSKVHDPI---PLRPKEI---I 656

Qy 504 SPLKILTCMQGHEIRROVQHI-EVDPDWEAAIAIQOLKNILLMPQEWG-----ACDE 557

Db 657 SLLTLCFLFNGAWKIKRKEGEHVLHEDQNFISLEYTTSIYSIIQTAEKVSKSDSIDS 716

Qy 558 ELLIWAYKECHKAVMRCSTSPISSSKTVVOSCGHSELTCKSVRSBDLVSIHLPLSRTLAG 617

Db 717 KLFANA-----IRIISFIGNSLTYKLIYDSEVIKFSVSHERVAFMPPLQTMLSF 768

Qy 618 LHVRLSRLGAVSRLLH---EFVSFEDQVEVLVEYPLRCLVLVAQVVAEMRRNGLSLSIQ 674

Db 769 LIEKVSLEKDAVEALEDCSDFLKSDFS-----LRSVLCSDQDVGFWNGMSVLHQ 820

Qy 675 VFYQDVKCREM--YDKDIIMLQIGA--SLMDPNKFLLLVLRVELAEAFNKTISTKQ 730

Db 821 ASY---KNNPELGSYSRDIHLNQLAILWERDDIPRIIYNILDRWELLDWFTGEVDYQHT 877

Qy 731 DLIKOYNTLIBEMLQVLIVIGERYVPGVNTYK---EVTREIHLICIEIPMHSATA 787

Db 878 VYEDKISFIIOPIAFIQILTERQYKTFSSLKDRMDQIKNSIINYLMFPLSYSKLL 937

Qy 788 KNLPEN-ENNETGLENVINKVATPKP-GVSGHGYVELKDESLKDFNMVYFYHYSKTQHSK 845

Db 938 RSPVDYLTEDTFEFDALAEVSVFVEPKGLADNGVFKLK-ASL-----YAKVDPK 987

Qy 846 AEHMQK-----KRRQENKDEALPPPPPEFCFPAFSKVINLLNCDIM---WYLIT 893

Db 988 LLNLNENEFESSATIIKSHLAKDKDBIAKVVLIPQVS-----IKQLDKDALNAGFTRNT 1041

Qy 894 VFERAI-----DTDSNLWTEGMLQMAFHILALG-LLEEKOOLKA-----PEEBVTFD 940

Db 1042 VPAKVYKLLQVCLMDEUSTFLNELHLVHGIFRDELLINGKDSIPEAVLSKFCINLLLS 1101

Qy 941 FYHKASRLGSSAM--NIQMLEKLKGIPOLEGQDKMITWILQMFDTVTKRLKSKCLIVAT 998

Db 1102 IANAKSDVFSESI VRKADYLLKEM-----IMKPNELFESLIA- 1139

Qy 999 TSGSSEIKNDEITHQEK-----AERKKAERAAHLRQKIMAOHSAQKNFIETH 1048

Db 1140 -----SFGQYVNDYDKKLRQGVNLQETEKERKRLAKHQARLAKFNNQOTKPMKH 1194

Qy 1049 KLMYNTSMPCKEBSIMBEESTPAVSYSRIALGPKRGSVTEKEVLTCILCOBQEVK 1108

Db 1195 ESEFD---EQDNDVDHV-----GEKYSESDFTCALCQDSS--- 1227

Qy 1109 IENNAWLSACVQKSTALTQHRGKFIELSGEALDP-----LFMDPDLA 1151

Db 1228 -STDPPVIFA-----YHDHSPIFRPGNIFNPNEFMPMDGFYNDDEKQAYIDDDVL 1277

Qy 1152 YG-TYTGSG-----HVHAAVCWKYFPAVQLSSQRIHVDLFDLESYELCLCK 1201

Db 1278 EALKENGSCGRKRVFVSCNHHIHNCKFKRYVQKGRPS-----SNAPICPLCQ 1324



QY 1202 SLCTVPIPIPLQOKINSENADALAQLLTLARWIOQTVLARIISGYNIRHAKGENPIPIFF 1261  
DB 1325 TFSNCTLPL--CQTSKANTG-----LSLDMFLESEL----- 1355  
QY 1262 NOQMGDSTLEFHSILSGVSESSIKYNSIKEMVILPATTIYRIGLKVPPDERPRVPMLT 1321  
DB 1356 -----DTLS--RLPFPFTENTRTINSISLMI----- 1381  
QY 1322 WSTCAFTIOAIENLLGDEGKPLGALONRONGKALKMQPAVAQRITCPOVLI-----Q 1375  
DB 1382 -SOCQGFDAKVRANRPNFHKVDSILSVHWANTISML---EIASRLEKPYISIFPRSRQ 1437  
QY 1376 KHLVRLLSVVLPIKSEDTCLASIDLPHVLGVAVLAPPSLYWDDPVDLOPSSVSSYNH 1435  
DB 1438 KY-----KTLKNI-----LVCIMLFTFVIGK-----PSMEF-EPYPOQPDVWNO-Q 1478  
QY 1436 LY-----LPHLITWAHMLQILLTVDGTPLAQVQEDSEASASFPABISQ----- 1482  
DB 1479 LFQYIVRSALFSPVSLRQVTREALTTFPSRQFLDRDFLOGLSDAQVTKLYAKASKIGDVLK 1538  
QY 1483 -----YTSGISGCDIPGWLWVLSKNGITPYLRCAALP---PHYLL 1520  
DB 1539 VSEQMLPALRTISDVMEGLDSESIIVDLAYTLKSL-----LPTIRRCVLFKVLHEL 1594  
QY 1521 GVTTPPELHTNSAGEVSALCSYLSLPTNLFLLFOEWTDTVRPQLQRCWADPALLNCLQ 1580  
DB 1595 KDSNETLVINGHEVEE-----LEFEDTAEFVNKALMKITEKESLVDLT 1641  
QY 1581 KNTVVRPRKN-----SLIELPDDYSCLLNQAASHFCRPSAD-----DERKHPV 1625  
DB 1642 QESIVSHPYLENIPYEVCGIILKIDLSKLVNTVTSQKEIKLEERSQHMKNADNRLDFK 1701  
QY 1626 LCLFCGAIL-----CSQNICCOEIVNGEVEGACIFALHCGAGVCIFLK 1669  
DB 1702 ICLTCGVKHLRADRHMTKLNKCFKPGAFMPNSE-----VCLHLT 1747  
QY 1670 IRECRVVEGKARGCAYPAPYIDYGETD-PGLKGNPNLHRSRYRKLHLVWQHCII 1728  
DB 1748 QPPSNIFI-----SAPYLNHSGVEGRNMRGDLTLNKRKYEHLNRLWINNEIP 1797  
QY 1729 EETARSQE-----TNQMLGCFN 1745  
DB 1798 GYISRVMGDEFRVTLISNGLFAFN 1822

RESULT 6  
T30554  
ubiquitin-protein ligase E3 - yeast (Kluyveromyces marxianus var. lactis)  
C:Species: Kluyveromyces marxianus var. lactis, Candida sphaerica  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T30554  
R:Waller, P.R.H.: Varshavsky, A.  
submitted to the EMBL Data Library, April 1998  
A:Description: Kluyveromyces lactis UBRI, the recognition component of the N-end rule pa  
A:Reference number: Z20856  
A:Accession: T30554  
A:Status: preliminary; translated from GB/EMBL/DBJ.  
A:Molecule type: DNA  
A:Residues: 1-1941 <NAL>  
A:Cross-references: UNIPROT:O60014; EMBL:AF061554; NID:g3114670; PID:g3114671; PIDN:AAC1  
C:Genetics:  
A:Gene: UBRI  
C:Superfamily: ubiquitin-protein ligase

Query Match 7.0%; Score 647.5; DB 2; Length 1941;  
Best Local Similarity 20.5%; Pred. No. 2.9e-31;  
Matches 409; Conservative 319; Mismatches 689; Indels 581; Gaps 91;

QY 35 FYTAFHLHLAQVPEIYFAEMDPDLKQESVQMSIPTPLEWYLFCEDPDLCLEKLKHS 94  
DB 51 YYYTMTSDSGLLPHHMTATNREBFKNVDQAMEIKLSKPMYKIDENG--HSPFNHAG 108  
QY 95 AFOLCGRVFKSGBTYTCRDCADPTCVLCMDCFQDSVHKNHRY--KMTSTGGGFCDCG 152

DB 109 --RIGAKFRVGEPIYRKECECSFDDTCVLCVNCFNPKDVGHHVYTSICTEFNNGICDCG 166  
QY 153 DTEAMKTGPFVNHPEGRAGTIKENSK-----CPLNEEVIYQARKIPPSVIKY 200  
DB 167 DKEAW-----NHELNCKGA-EDNGRLEDEPDHDGKISKMLESVLIELPDPHIDVFNQ 218  
QY 201 VVE--WTIWE-----EKEKLPELQIREK-----NER----- 225  
DB 219 NIEPLTTIOKPLIAKLYRNFPEREYEQADMRLRLAYRNQYMDDEESSNKRHLTSLDPLS 278  
QY 226 ---YCVLPNDHEHSHYVIVISLQALDCELABAOLHTTAIDKEGRRAKACA-VAACQ 280  
DB 279 TLKDYAILYYDEFNYSQASAAIROG-GPDNKHIDLLTAKIDSEGRSLRCSADIASLM 337  
QY 281 EAKEDIKSH--SENVSQHPLHVEHSEMAHOKFALRLGSGWMNKIMSY--SSDFROIIFCQ 337  
DB 338 GRIFSVQSNGLSCTITQ---WYELHQEAC---KYSI---MWINDCLINPNSITFQSLFRN 388  
QY 338 A-----CLREEDP-----SENPCLI----- 352  
DB 389 AIGKVLCSKVEPFYQSIDMTSVVDRYFSDSYLSDDPYLVADHSVLGEGVKIPLGRHKSLD 448  
QY 353 -----SRL-----MLWDKLYKGARKILHELIFSSFFMEMEYK 385  
DB 449 PGDISAISPILANKVIAEDHHEYNLSRLOYVLFLENRYWKGKRIKIVQDLIPTLASSAVQK 508  
QY 386 KLPAMEFVKYKQLOKEYISDDHDSISITALS---VOMPTVPTLARHLEE-----Q 435  
DB 509 PHMTDQVIEIPHMTRS--GTFMDREPQUTLSRESVQLTCTPTTAYTSIFHSGHFNLIW 566  
QY 436 NVISVITE--TLLE--VLPEYLRNNKFNFGYSODKLGRRVYAVICDLKYI-----LISKP 487  
DB 567 SVIDVPDFTMDEGTLVWQVRQSRNP--SKYSISFKQGLYAVETLSSKITDPNLLAKP 624  
QY 488 TIWTERLRMQFLEGFRSFLKILT---CMQGMEEI--RRQVQHI--EYDPDWEAAIAIQMQL 542  
DB 625 -----GEFIMVTLCFLFNGAWKIKKQREGEHVLREDQHFIPYLEYTTISV 669  
QY 543 KNILLMF-----QEWACADEILLVAYKECHKAVMRCSTSFSSSTVTVQSGH----- 591  
DB 670 YSIIQTDFKVLQOSKDHIDQRLIGA-----INLDSFL-----GHRNLSYK 711  
QY 592 ---SLETQSYRVSEDLVSIHLPLSRTLAGL--HVRLS--RLGAVSRHLHFVSEDFQVEVL 645  
DB 712 LYKDPELIKFOISKQSVSPMNPVHTLFCFLVQHVPLQVSIQVLSQSKDYLIVISDF----- 766  
QY 646 VEYPLRCLVLAQVVAEMRRNGLSISQVFTYQDVKCREEM--YDKDIIMLQIGASLMD 703  
DB 767 ---ALRSVVLCSQIDIGFVWRNGMSVLHOSAYY---KNNPEMSSYSRDIQLANQL-AFLIE 819  
QY 704 PNKELLV--LQRYELAEAFNKTISTKDDDLIKQYNTLIEEMQLVLIYVIGERYVPGVG 760  
DB 820 KNDFORVIYNMLDRWELDFWFGSVSTETVYDDKISSIIQQFVAPLYQLLVER----- 873  
QY 761 NVTKEBVTMRE-----IIHLACIEPMHSAKALNPEN--ENNETGLENVINKVATP 810  
DB 874 DPKKFDLLEETQLYNKNAIYKLYAEPLSYTDLLNDIPDYLTVESVSQFDVLEEVSY 933  
QY 811 KXPGVSGHGVYBELKDBESLKD-----FNM--YFYHYSTQHKAKAEMQKRRKQENKDBA 862  
DB 934 IEPKGSKIWCFTKQCTLQENRFLNNGNDFEHSATIVKSHLADSKEKRAKIIKVPQ- 992  
QY 863 LPPPPPEFCPAFSKVINLLNCDIMMYILRTVFERAIDTDSNLTWTEGMLQMAFHILALGL 922  
DB 993 -----LLELDELDPGAR---ELGSPFRTNLFA---KLIFKLLKLA 1027  
QY 923 LEEKQLOKAPREEVTFDFYHKASRLGSSAMNIQMLEKLGIPOLEGOKDMITWLOMF 982  
DB 1028 -----SDSSFSFTYELLHLIHAIFRDEENVNCKDSLPEAYISKPCIDLLSIV 1075  
QY 983 DTWKLREKSCILIVATTS-----GSEINKNDRIITH-----DXE 1015  
DB 1028 -----SDSSFSFTYELLHLIHAIFRDEENVNCKDSLPEAYISKPCIDLLSIV 1075  
QY 983 DTWKLREKSCILIVATTS-----GSEINKNDRIITH-----DXE 1015







QY 948 -----LGSAMNIOMLEK-----LKGIPOLEGOKMIT-----WILQMF 982  
Db 1279 LIKPREMETETLVCAFIRLLTKETRKNGALVVVFKGILSGEYDKBKVTGCKMYLARFV 1338  
QY 983 DTVKRLREKSLIVATTSGESIKNDIBTHDKKAE-----RKRKAAEAHLRQKIMAQ 1036  
Db 1339 TILTSLFVARQIIEGKKEELRISKHSRQEKQKAPMDPVKAAKAEAKRMEALM-- 1396  
QY 1037 MSALQKNPIETHKLMYDN--TSEMPGKEDSIMBEESTPAVSRIALGPKRGPSTVEK- 1093  
Db 1397 QNSAKKSQAKMKTKEGMDTDAENVKVDPSQQRKV-----YECPICGEQNAPNTVENP 1451  
QY 1094 -----EVLITCILCOEBOEVKIENNAVL-----SACVQKSTALTQHRGKPIELSGEA 1140  
Db 1452 FGLAKLSTNPICEEQIDASINTIDILLKFDYEYHVSANLQSETRRFRFFSKRQATFEN 1511  
QY 1141 LDPLFMDPLAYGTYTSGCGHVMHVCWQKY-----PRAVOLSSQORIHVDLPD----- 1189  
Db 1512 ODIVKVNPLV-GTDLTKCHAAHACFNAYRASLVGSFOEKKLENEK-----KYDGHQ 1566  
QY 1190 -LESGEYLCPLCKSLCNVTPI-----IPLOPKINSENADALAQALLTLARWIQTVLAR 1242  
Db 1567 STDRREVGCPWCRTYVNVIVPMSFDKPYTPIKTPASPMYSDV-----N-----R 1611  
QY 1243 ISGYNIRHAKENPIPIFFNQMGDSTLEPHSILSPGVSSIKYN--SIKEMVILFATT 1300  
Db 1612 VMDVVLKKAAG---PVFQED-----ERNVKYSTNYSTREGGLFELY 1650  
QY 1301 IYRI-----GLKVPDERDPRVPM-----L 1320  
Db 1651 IGRHSADLAERKSSQORCTVSTLMVSLAVVMVVERSSIMRKMGPERRKNQRNMTHELM 1710  
QY 1321 TWSTCA-----FTIOAIENLGDGKPLFGALQNRQNGKALMQFAVAORITCPOVL 1373  
Db 1711 TASVATSKVDVFDVALSALTNLFAKTENSFSRPPSEQSTSEPAQCTEAPN---PEAV 1767  
QY 1374 IQKHLVRLLSVLPNTKSEDTPCLLSIDLPHV-----LVGAVLAPPSLYWDDPDVLO 1425  
Db 1768 VGLSSDEMAAMITKPLRKOTPDCLSKPLFALDPKATLVRLAVLIDNQSLTKDIQREIA 1827  
QY 1426 PSSVSSSYNHLYLPHLITMAHMLQILLT-----VDGLPLAQ-----VQDSERASHA 1473  
Db 1828 QNMIFASLGVWSSYTLTCL-----ILRTGEEKISALNKGEPKIQGLSDHLQSEAEACQA 1882  
QY 1474 ----SFFAEISQVTSIGCDIPGWVLMVSLKNGITPYLRCA-ALPHYLLGVTP-PEE 1527  
Db 1883 LTYNTYFPHLAQRLB-SPDVEPTDEYICKTMNSCLIEFLRFYSLELPHCNLGFNDVNNQ 1941  
QY 1528 LHTNSAEGE-----YSALCSYLSLPTNLFLLFOEYWD-----TVRPLLQRCADPALL 1575  
Db 1942 IHSRNDLDNLVRLVGINAQNVPGK-----ANYWTRAFRQLTWMAKLR-IPQPCIV 1995  
QY 1576 NCLKQKQNTVVRPKRNSLIEIPDDYSLINQASHPCPRSDADDERKHVLCFLCGAILC 1635  
Db 1996 EPLAWK-----PRR---ILKPTNPDELFGYFHREC-NKSSVPSYPMVCLFCGEILC 2045  
QY 1636 SONICCO-----EIVNGEYVACIFHALHCGAGVICILKIRECRVVLVEGARGCA 1686  
Db 2046 -LNDCCRMHQHSGSDRVISMSEIEA---HABDCSSSGLFSLVSTSSVMVVSRGK-QAAI 2100  
QY 1687 YPAPLYDEYGETDPLGKGNPLHLRSRYRKLHLVWQOHCITIEEARSOE-----TNOMLF 1742  
Db 2101 WGTIYLDAAKEDRNLRKPKPLFLCESRLKWLLEYDAE-----QEWQRPQWFMNTNSQAF 2156  
QY 1743 GFN 1745  
Db 2157 TSN 2159

RESULT 10  
T48252  
eceriferum3 (CER3) - Arabidopsis thaliana

N;Alternate names: protein T1E22.70  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C;Accession: T48252  
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft  
submitted to the Protein Sequence Database, March 2000  
A;Reference number: Z24489  
A;Accession: T48252  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-795 <REV>  
A;Cross-references: UNIPROT:Q9LZ94; EMBL:AL162874  
A;Experimental source: cultivar Columbia; EAC clone T1E22  
C;Genetics:  
A;Map position: 5  
A;Introns: 203/2; 220/3; 428/1; 488/3; 666/2; 701/2; 723/3; 749/3; 772/3  
A;Note: T1E22.70

Query Match 3.7%; Score 343; DB 2; Length 795;

Best Local Similarity 23.0%; Pred. No. 4.6e-13;

Matches 190; Conservative 123; Mismatches 274; Indels 240; Gaps 43;

QY 1033 INAAQS-----ALQKPIETHKLMY-----DNTSEMPGKEDSIMBEESTPAVDSRTA 1081  
Db 56 LMLQSLDDTISANMIESIKARLIGNGQTEKRSSDGRGKDESME---SLEIAMYQTVR 112

QY 1082 LGPKRGPSTVEKVELTCTILCOEBOEVKIENNAVLACVOKSTALTQHRGKPIELSG--- 1138  
Db 113 -----NKIENWVQSLTRVDHQHEAEN-----CSEKNSV-----GGSTLOGRRPP 153

QY 1139 -----BALDPLFMDPLAYGTYTSGCGHVMHVCWQKYFEAVQLSSQQR 1183  
Db 154 DIRSRQTSRRPDAGSDGPHPIDCD-----GVYLSGCGHVAHQSLERLYLSKERSGR 208

QY 1184 HYD-----LPDESGEYLCPLCKSLCNVTPIIP-----LOPQ-KINSENADALAQALLT 1231  
Db 209 VFEGAHIVDLKCKBLCPVCRRLANSVLPCFCDLCSVKLQDSPRTKLRRKDALQPSL- 267

QY 1232 LARWITQTVLARISGVNIRHA-----KGENPIPIFFNQMGDSTLEPHSILSPGVSS 1283  
Db 268 ---WLSEALCL-----LRSAAEVIIDGDRGKTVP-----QGDGPRR----- 301

QY 1284 IKYNSIKEMVILF-----ATTIYRIGLKVPPDERDRPVMLTWTSTCAFTIOAENL 1337  
Db 302 -KDLASVSKMLWDFYPPKPEDKTKLKLWL--PPQS-----IWMWDTLYKSLISME--IG 350

QY 1338 DEG-----KPLFGALQNRQNGKALMQFAVAQR-----ITCPQVLI-QKHLVRL 1381  
Db 351 TRFAKNSMLPVYCIDSLVEELKTSKGTILSVLLRVVQSRTKNTIHRQRFVGMKHLAES 410

QY 1382 LSVVLPNLIKSED-----TPCLLSIDL-----PHVIVGAV--LAPPSLYW 1418  
Db 411 ICYGVSSSSSSSIFGSEGTGSLKNIDLLMNRASDPVLAHDPFSSLMALFCLPPFPLTC 470

QY 1419 DDPVDLPQSPVSSSYNHLYLPHLITMAHMLQILLTVDGLPLAQVQEOSEEAHSASSPPA 1478  
Db 471 EESL-----LSLVHIIHVSLSVLTQTVIAYACRP-----SELSELNFENL 512

QY 1479 EISQVTSGS-----IGCDIPGWVLMVSLKNGITPYLRCAALFFHYLLGVTPPE 1526  
Db 513 DISNALRESGWEVPRSNMMDLSCDIKD-----TIKYSYLPFLRCALLKWL--KSTPR 565

QY 1527 ELHTNSAEGEYALCSYLSLPT--NLFLLP--QEWDTVRLQRCADP--ALLNC-LK 1579  
Db 566 KLH-----EESDMFPLSDPTTDNDMDFYSPQSELNHHVQELKMFNIPPIIDILNDELL 619

QY 1580 QKNTVW-----RYPRKENS-----LIELPDDYSLINQASHPCPRSDADE 1620  
Db 620 RSSTQIWLQHQFQREYRNVNRKRSICITPVFPQLMKGNLYQDLLQRCIKRCV-NCTKV 678

QY 1621 RKHPVLCFLCGAILCSQ--NITCCOBIVNGEYVACIFHALHCGAGVICIFLKIRECRVVLV 1678  
Db 679 IEEPVLCLCGS-LCSPIWSPCCRE-----SGCPNHATCGAGTVGLLIRR--TTILL 729

QY 1679 EKGRCAYPAPVLDYGETDCLKRGNPLHLSSRRYKXHLVWOQH 1725  
DB 730 QRFARQSPWSPYLDFTGEBEDIMIRGKRLYLNEERYAALTIVGVSH 776

RESULT 11  
T40238  
hypothetical protein SPBC32F12.14 - fission yeast (Schizosaccharomyces pombe) (fragment)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40238  
R:Moreno, S.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, June 1998  
A:Reference number: Z21915  
A:Accession: T40238  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-271 <MOR>  
A:Cross-references: EMBL:AL023796; PIDN:CAAL19375.1; GSPDB:GN00067; SPDB:SPBC32F12.14  
A:Experimental source: strain 972h; cosmid c32F12  
C:Genetics:  
A:Gene: SPDB:SPBC32F12.14  
A:Map position: 2

Query Match 2.4%; Score 218; DB 2; Length 271;  
Best Local Similarity 26.0%; Pred. No. 4.7e-06;  
Matches 60; Conservative 34; Mismatches 97; Indels 40; Gaps 6;

QY 60 EKQEEVSQMSIPTLEWY-----LFGDDPDICL-----EKLXHGSAFQJC 99  
DB 37 ESNKSLNNEVSALLGYDHTLWNTLLPERPIDASFLRRACQHGSEGBYRHGTCESKC 96

QY 100 GRVFKSGEYTCRCDAIDPTCVLCMDCFQDSVHKHRYKMTSTG-GGFCDCGDTFAWK 158  
DB 97 GHIFRGEVFRKTCSDVNSALCVKCFRATSHKHDSFTVSAGSGGCCDCGNAAMI 156

QY 159 TGPFCVNHFRAGTIKENSRCPLENEVIVQARKIPFSPVKKVVENTIWEER--KELPPE 216  
DB 157 GDVSKCIHSHBEDATISNDMIDPEIKLNSIQTTIDCVLDFVDFVSCPENLKMPTL 216

QY 217 LQI--REKNER-----YYCVLFNDEHHSYDHYIYSIQRALD 250  
DB 217 ESILQDEKTSRSENKYGDIDDSNMYSLVNLWDEKHSFKQFEQITTALE 267

RESULT 12  
JCS5837  
364K Golgi complex-associated protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C:Accession: JCS5837  
R:Toki, C.; Fujiwara, T.; Sohma, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.  
Cell Struct. Funct. 22, 565-577, 1997  
A:Title: Identification and characterization of rat 364-kDa Golgi-associated protein rec  
A:Reference number: JCS5837; MUID:980933490; PMID:9431462  
A:Accession: JCS5837  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-3187 <TOK>  
A:Cross-references: UNIPROT:Q63714; DBJ:D25543; NID:g516825; PIDN:BA05026.1; PID:g5168  
C:Comment: This protein plays a role in the formation and maintenance of the characteris  
C:Superfamily: giantin  
F:49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status predic  
F:3165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match 2.2%; Score 205.5; DB 2; Length 3187;  
Best Local Similarity 20.2%; Pred. No. 0.0011;  
Matches 243; Conservative 182; Mismatches 472; Indels 305; Gaps 49;

QY 125 MDCFQDSVHKHRYKMTSTGGFCDCGDTFAWKTPFCVNHFPBRAGTIK-----ENS 178  
DB 1057 IDLQBEITENQATIQKFIPTGTDAGDGSVAKETSV----SSPPRAGGGEHWKPELEGK 1112

QY 179 RCPLNEEVIVQARKIPFSPVI-KYVVEMTIWBEKEKLPPELOIREKNERYYCVLFNDEHHS 237  
DB 1113 IVDLEKEKTQLOKLOEALISRKAILKCAQKEKHLKEELKEQDAYRHLQEQFQGSKE 1172

QY 238 YDHYIYSIQRALDCELAQAQLHTTAIDKXGRRAVAGAYAAQCAQEKEDIKSHSENVQ-- 295  
DB 1173 NENIRAPLRQLQAKESTDQQLPGT-----GQEPHTGSEGLSECTEPASBSDLHAAQPS 1227

QY 296 HPLHVEVHLHSEIMAHQ-----KPALALGSMWKMIMSYSDFFQIFQACILREBP 344  
DB 1228 HPGETATTQATVSVAAQIQDLKBEIVBEKELEL-----KISSTTSBELTKSEVILLQEQ 1282

QY 345 DSENFCLISRLMLWDKLYKGARKILHELIFSSP-FMEVEYKCLFAMFVVKYKQLOKEY 403  
DB 1293 INEQGLEIQNLKAASHEAKAHEQKQLESOLKIALDHLKTLQPEL-----ETLQKHV 1338

QY 404 ISDDHRSISITALSVQMFVPTLARHLIEBQNVISVITETLLLEVLPRYLDNRNKNFQG 463  
DB 1339 GQKEEVSYLQGLGKEKQTLTTVQTEM-EEQ-----ERLIKALHTQLEMAQKEH--- 1387

QY 464 YSQDKLGRVAVICDLK-----YILISKPTTWTERLMQFLGFRSFLKILTCMQGMBEI 518  
DB 1388 --EERLKQVQVEICELKQPKLEBESAKQOQLQKLAALISRKALKENKSLQEQLS 1445

QY 519 RQVQGHLEVD-PDWEAAIAIQMOLKNILL-----MFOEWCADEBELLVAYKECHKAVNR 573  
DB 1446 ARDAVEHTKSLADVESQVSNQKDALGLKALLQE-----ERDKLIV---EMDKSLLE 1498

QY 574 CSTSPISSTKVQVCGHSLKTSYRVSDELVSIH---LPLSRTLALHVLRLSR-LGAVS 629  
DB 1499 -NQSLGGSCSKLALGGLTEDKB-KLMKELESVRCSTABSTEWQEKHKLQKEYEVILL 1556

QY 630 RLHFVFSFEDFQVFLVVEYPLCLVLVAQVVAEMW-----RNGLSLSIQVFPY 677  
DB 1557 QSYENVNSAEERIQHVVE-----SVRQEQEVYAKLRSAESDKREKQLODAEQEME 1609

QY 678 YQDVKCREEMVDKDIIMLQI-----GA-----SLMDPNKFLLLVLQRYEL 717  
DB 1610 EMKEKRRFPKSKQKQKILEBENDRLRAEAPVCGANESMALISSNASLKEELERITL 1669

QY 718 ABAPNKTISTKQODLHKQYNTLIEMLQVLIYVGERYVPGVGVNVTKEBVTWREIHLIC 777  
DB 1670 E---YKTLSEKPEALMAEKNTLSEETRNMLKQVEAQELKQASLETTEKSDPKDV----- 1721

QY 778 IEPMPHSAIAKULPENENNETG-LENV-INKVATFKPGVSGHGVYELKDESIDKDPNMYF 835  
DB 1722 IBEVTEAVVGVKSQEQDSLENAKLEDAEATLLANSAPGVSE-----TFSSHDDINNYL 1775

QY 836 YHYSKTQHSKAR-HMQKKRRKQ-----ENKDEALPPPPPPPEPCAPFSKV-----IN 880  
DB 1776 QQLDQLKGRIAELEMEKQKDBLSQTLNEKNAL-----LTQISAKDSSELKLEEEVAKIN 1831

QY 881 LLNCDIMMYILRT--VFERRAIDTDSNLMTEGMLQMAFHILALG----- 921  
DB 1832 MLNQIQELSRVTKLKTABEEKDLEERLMLNQLAELNGSIGNYQDVTDQAIKNEQLE 1891

QY 922 -----LLEEKQOLQKAP---EEVTFDFYHK-----ASRLGSSAMNIQML 958  
DB 1892 SEMONLKRCSVSLBEKQOLVKEKTKVSEIRKVEYMEKIQGAQKGPQGIHAKELQ-ELL 1950

QY 959 LSKLXGIPQLEGQKDMITW---ILQMFDTVVKEL----- 988  
DB 1951 KEKQEVKQL--QKQIRYLGRISALEKTVKALEFVHTBSQKDLATKGNLAQAVEHHK 2008

QY 989 -----REKSLIVATV-----SGSESIKND-----EI 1010  
DB 2009 AQAEELSSFKILLDDTQSEBAARVLADNLKLKELQSNKESIKQIKQKDBDLRLRLEQASE 2068

QY 1011 TIDKEKAERKRAEAAARLHRQIKMAQ-----MSALQKNFIETHKLMYDN 1054  
DB 2069 KIRKEKKNQOEKLDLA--LHREKAHVEDTLAEIQVSLTRKDKDKMKELQQSLDSTLAQLAAP 2126



QY 1055 TSEMPGKEDSIMEESTPAVDYSRIALGPKR-----GPSVTEKEVLTLCILCOBEQEVKI 1109  
Db 2127 TKSMSSLDQ-----DNRVIDEAKWEQRFDAIQTK-----BEVRLKE 2166  
QY 1110 EN 1111  
Db 2167 EN 2168  
RESULT 13  
A56539  
giantin - human  
N:Alternate names: macrogolin  
C:Species: Homo sapiens (man)  
C:Date: 19-Oct-1995 #sequence\_revision 26-Jan-1996 #text\_change 09-Jul-2004  
C:Accession: A56539; S37536  
R:Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M.  
Mol. Cell. Biol. 14, 2564-2576, 1994  
A:Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein  
A:Reference number: A56539; MUID:94187728; PMID:7511208  
A:Accession: A56539  
A:Molecule type: mRNA  
A:Residues: 1-3259 <SES>  
A:Cross-references: UNIPROT:Q14789; EMBL:X75304; NID:g405714; PIDN:CAA53052.1; PID:g405714  
C:Genetics:  
A:Gene: GOLGB1; GCP; GCP371  
A:Cross-references: GDB:454958  
A:Map position: 3q13.31-3q13.31  
C:Superfamily: giantin  
C:Keywords: coiled coil, Golgi apparatus, transmembrane protein  
F:3238-3254/Domain: transmembrane #status predicted <TMN>  
Query Match 2.1%; Score 190.5; DB 1; Length 3259;  
Best Local Similarity 18.0%; Pred. No. 0.0096;  
Matches 221; Conservative 207; Mismatches 424; Indels 377; Gaps 49;  
QY 164 VNHEPRAGTIKENSRCPLNEVIVQARKIFPSVIVKVVEMTWEEKEKLPBQL--IRB 221  
Db 12 VLHLGGDDTDQNMRAPLDPELHQBESDMFNNTTQEDVQERLAYAE-QLVVELKDIIHQ 70  
QY 222 KNERYYCVLFNDEHSHYDVIVSYLQALDCELAEAQLHT-----TATDK--EGRAVKAGA 275  
Db 71 KOVQL-----QKDEALQBERKAANKKUKLHAKAKLTSINKYIBEKAAQGGTV 121  
QY 276 YAAQCBAKEDIKSHSENVSOHPLHVLHSEIMAHQKPAIRLGSWMNKIMSYSSDPQRI 335  
Db 122 LPTEPQSEQLSKHDKSSTEEEMIEKIKHLQKEBELISTL----- 163  
QY 336 CQACLREPPDSENPCLISRLMLWDKLYGARKILHELIFSSFFMEYKFLPMEFVKY 395  
Db 164 -QAQL-TOAQAEQPAQSS-----TEME-----EFVNM 188  
QY 396 YKLOKEYISDDHRSISITALSVQMFTPTLARHLIEQN-----VLSVIT 442  
Db 189 KOOLQKB-----EEFISTLQQLSQTQAEQAQVVRKEDARFETQVRLHEDLLQVLT 242  
QY 443 ETLEVLPEYLDNRNKNFQGSQDLGRVAVICDKYLKILS-----KPTIWTIERLM-- 496  
Db 243 QADVETEMQKRLVLRQLKEEESLVGR--AQVDLLQQLTAAQRNQLLSQQLQOME 300  
QY 497 -----QPLEGRSPFLKILTCQGMETRRQVGHIVDPDWEAAIAIQMLKNILLMP 549  
Db 301 AEHTLNRNTVETREESKILKEKLEVAERKLSFH-----NLQEMHLLQF 349  
QY 550 QEWACDDEL--LLVAYKECHAVMRCSTFI--SSKT--VVQSCGHSLETYSYRVSDEL 604  
Db 350 EQAGQAQAELEGRYSALQKHAEMBEKTSHLISLQKTQGLQSLQACDALKQNSKLLQDK 409  
QY 605 VSIHLPLSLTLAGLHVLSRLGA-----VSRL-----HEF-----VSPEDFQVLEVLVEPLRC 652  
Db 410 NEQAVQSACTIQLELDQLQKSKESISQFLNRLPLQHQHETASQTSFPDVTNEGTAVENTE 469  
QY 653 LVLVAQVVAEMWRRNGLSLISQVFYYQDVKCREMYDKDIIMLQI-----GAS 700

Db 470 IASLQRRVVELENKRGALLSSI--ELBELKAENEKLSQITLLEAQNRTEADREYSEIS 528  
QY 701 LMD-PNK-----FLILVLQRYELAEA 720  
Db 529 IVDIANKRSSAABESQDVLENTFSSQKHKELSVLLEKMAQBEIAFLKQLQKQKAEAA 588  
QY 721 FNTISTKQDILIKY--NTLIBEMLOVLIYIVGRY----- 755  
Db 589 DHEVL---DQKEMKQMEGHGIAPIKMKVFLDQDQFPLMPNEBESSLPAVEKQAEQASTHQ 645  
QY 756 -----VPGVGVTKERTVTMREIHIHLLCIEPMPHSA-- 785  
Db 646 SRTSEISLNDAGVELKSTKQDGDKSLSAVPOIGQCHQDELB--RLKSQILELNLPHKAQ 704  
QY 786 --IAKLPENNETGLENVINKVATFKPGVSGHGV-----ELKDESL---KDFNNYF-- 835  
Db 705 EYIEKVLDEKAEISNLNLQIEB---FKGNADNNNSAFTALSEERDQLLSQVKELSMVTE 761  
QY 836 --YHYSKTQHSKAEHMQKGRKOENK---DEALPPPPPPFCFAPSKVINLLNCDIMMYI 890  
Db 762 LRAQVKQLEMLNLAEAERQRLDYESTQAHNLLTE-----QIHSLSTEAKSKDVKIEV 814  
QY 891 LRTVTP-----RAIDTDSNLWTEGMLQWPH 916  
Db 815 LQNELDDVOLQFSRQSTLIRLSQSLQNKESVLEGAERVRHISSEKVELSQALSQKLE 874  
QY 917 ILALG--LLEEK---QOLQKAPEE-----EYTFQFHKASRLG---SSANNIOMLEK 961  
Db 875 IYTMQLLLEKRDVETLQOTIEEKDQVTEISFSMTKQVQNEKFKSLGVIEIKTLKQ 934  
QY 962 LKGI-----POLEGKDMITWILQMDPTVKRLREKSLIVATTSSESINKNDEITHD 1013  
Db 935 LNLSSRAEAKKEQVEEDNEVSSGLKQNYDEM-----SPAGQISKEELQHE 980  
QY 1014 -----KEKAERKKAERAAHLRQKMAQMSALQKQPIETHKLMYDNTSEMPGKEDSIMEE 1068  
Db 981 FOLLKKEKQRKQLQALINKKELLQVRSLER---ELANLKDESKKPISETERGEV 1037  
QY 1069 ESTPAVDYSRIALGPKGPSVTEKEVLTLCILCOB-----EOEYKIENNAWVLS- 1117  
Db 1038 EEDKENKEYSE-----KCVTSK-----CQIEIYLVKQITSEKVELQHIRKDLSE 1082  
QY 1118 --ACVQKSTALQHRKPIELSGEALDPL 1144  
Db 1083 KLAEEQFOALVKQMNTLQDKTNQIDLL 1111  
RESULT 14  
A57013  
early endosome antigen 1 - human  
N:Alternate names: endosome-associated protein  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: A57013; S44243  
R:Mu, F.T.; Callaghan, J.M.; Steele-Mortimer, O.; Stenmark, H.; Parton, R.G.; Campbell, J. Biol. Chem. 270, 13503-13511, 1995  
A:Title: EA1, an early endosome-associated protein. EA1 is a conserved alpha-helical  
A:Reference number: A57013; MUID:95286647; PMID:7768953  
A:Accession: A57013  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1410 <RES>  
A:Cross-references: UNIPROT:Q15075; GB:I40157; NID:g1016367; PIDN:AAA79121.1; PID:g1016367  
R:Seelig, H.P.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S44243  
A:Accession: S44243  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-254, 'C', 256-257, 'LO', 260-276, 'A', 278-283, 'A', 285-519, 'D', 521-574, 'EQ', 577-  
A:Cross-references: EMBL:X78998; NID:g475933; PIDN:CAA55632.1; PID:g475934  
C:Genetics:



699	QY	-----ASLMDPNKFLLVLLQRYELA-----EAFNKTISTKQODLKIQYN-----	737
648	DB	ABAEKREALWEKNTHLEAQLQKAEAGAEAGLQADLFDIQESKEEIQKLSHSHQOEAAAT	707
738	QY	TLIEEMLQ-----VLIVIGERYVPGVGNVTKEBVTMREIITHLICIE-----PMP	782
708	DB	TQLEQLHQEAKQEVEVLARAVQKEA-----LVREKAALAE--VRLQAVZERDQLAAQLQ	760
783	QY	HSAIAKNLPENENNETGLEN-VIN-----KVATPKPGVSGHGVIYBLKDSLKDF	831
761	DB	GLSSAKELLESSLFEAQQNSVIDBPQGLEVIQIQTVTOAKEVIQGE--VRCLKLE----	814
832	QY	NMVFYHSYKTOHKSKEAHMKKRBKORNKDEALPPPPPPFCFAPSKVINLLNCIDMMYIL	891
815	DB	-----LDTESQABOEERDAARQAQAQBQKGTALBQOKAHEKVNQ-----L	858
892	QY	RTVFERAIDTDSNLMTGMLQMAFHIALGLLLEKQLOKAPPEEVTDFYHKASRLGSS	951
859	DB	REKWEK-----ERSWHQQLAKALESLEBKMELEMLRKEQQTMEAEIAQAREEBERTOAE	913
952	QY	AMNIOMLLEKLAGIIPOLEQCKDMITWILQMFDTVVKLRKESCLI--VATTSSSESIKNDE	1009
914	DB	SALCOM-----QLETEKERVLSULETLLQTKELADASQOELRLDRQMKVQKLKEQE	964
1010	QY	IT-----HDKBEKAERKKAABAAHLRQKWA-----OMSALOKNFTIETHKLMYDNTSEMPGKE	1062
965	DB	TTGIIQTQLOEAORELK-EAARQHRRDDLAALQEBSSSLQDKMDLQKVEDDLKSQLVAGD	1023
1063	QY	DS--IWE-----RESTPAVSDYSRIALPGKPGSPVTEKEVLTCLIQOEQBQVKIENNAWVLS	1117
1024	DB	DSQRLVEQVBQKLRQETQYNNRIQELER-----EKASITLSLMEKEQRLVLVLEAD--S	1076
1118	QY	ACVOKSTALTOH-----RGKPIELSGE-----ALDPLFMDPDLAYGTTGSCGHV	1162
1077	DB	ITQOELSAURQDMQBAQGEQKELSQAQMEILLROEVKEKEADFLAQEAQLLEEL----	1133
1163	QY	-----MHAVCWQKYFEAVQLSSQORIHVDLFDLESGEYLCPLCKSLCNTVIPITPIQPOKRI	1218
1134	DB	TEQQLRASLWAEAKAAQL-----HLRLRSTES-----QLEALAAEQ	1171
1219	QY	NSENADALAQLLTLARWITQVLARI	1243
1172	DB	PGNOQAQAQALASLYSALQOALGVS	1196

**THIS PAGE BLANK (USPTO)**